1982

Cooperative rule-learning: a metacognitive interpretation

Timothy Arthur Bender
Iowa State University

8224326

Bender, Timothy Arthur

COOPERATIVE RULE-LEARNING: A METACOGNITIVE INTERPRETATION

Iowa State University PH.D. 1982

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Cooperative rule-learning:

A metacognitive interpretation

by

Timothy Arthur Bender

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Psychology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Members of the Committee:

Signature was redacted for privacy.

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

	Page
COOPERATIVE RULE-LEARNING: A METACOGNITIVE INTERPRETATION	1
RULE-LEARNING AS PROBLEM-SOLVING	4
Rule-Learning Perspectives	7
Hypothesis-Testing and Rule-Learning	11
Metacognition	16
Model of cognitive monitoring	18
Improving metacognitive control	20
Vocalization of reasoning	21
Summary	25
FACTORS WHICH INFLUENCE RULE-LEARNING	26
Task Variables	26
Context characteristics	27
Paradigms	27
Type of rule	27
Pretraining	28
Group size	29
Cooperation	31
Ability groups	32
Summary	33
Stimulus characteristics	33
Array characteristics	33
Redundant information	34
Negative instances	35
Summary	37

	Page
Memory characteristics	38
Feedback characteristics	39
Person Variables	40
General mental ability	41
Personality variables	42
PURPOSE	44
EXPERIMENT ONE	45
Research Hypotheses	47
Method	48
Subjects	48
Apparatus	48
Procedure	48
Design	50
Scoring procedures	51
Results	52
Learning session	52
Transfer session	54
Transfer across sessions	57
Discussion	59
Transfer on successive problems	62
Transfer across sessions	63
Conclusion	64

	Page
EXPERIMENT TWO	66
Research Hypotheses	66
Method	67
Subjects	67
Apparatus	67
Procedure	68
Design	70
Results	71
Learning session	72
Transfer session	74
Transfer across sessions	75
Discussion	76
Transfer on successive problems	77
Transfer across sessions	78
Conclusion	78
EXPERIMENT THREE	79
Research Hypotheses	80
Method	81
Subjects	81
Apparatus	81
Procedure	81
Design	83

Page

Results 84

Learning session 84

Transfer session 89

Transfer across sessions 95

Discussion 108

Transfer on successive problems 109

Transfer across sessions 110

Conclusion 112

GENERAL DISCUSSION 115

Difficulty of the Arrays 115

Group Size 116

Vocalization 117

Problems 118

Conclusion 119

REFERENCES 121

APPENDIX A. SAMPLE INSTRUCTIONS 126

APPENDIX B. ANOVA SUMMARY TABLES 128

APPENDIX C. THREE- AND FOUR-WAY INTERACTIONS 153

LIST OF FIGURES

		Page
Figure 1.	Design of Experiment 1	50
Figure 2.	Mean Untenable Hypotheses for Group Size x Session in Experiment 1	58
Figure 3.	Mean Strategy Efficiencies for Group Size x Session in Experiment 1	60
Figure 4.	Mean Decision Times for Group Size x Sessions in Experiment 1	61
Figure 5.	Design of Experiment 2	71
Figure 6.	Design of Experiment 3	83
Figure 7.	Mean Trials to Solution for Group Size x Session in Experiment 3	97
Figure 8.	Mean Untenable Hypotheses for Group Size x Session in Experiment 3	99
Figure 9.	Mean Strategy Efficiencies for Group Size x Session in Experiment 3	101
Figure 10.	Mean Strategy Efficiencies for Group Size x Vocalization x Session in Experiment 3	103
Figure 11.	Mean Decision Times for Group Size x Session in Experiment 3	105
Figure 12.	Mean Decision Times for Vocalization x Session in Experiment 3	106

LIST OF TABLES

		Page
Table 1.	Means of all dependent measures for problems in the learning session of Experiment 1	53
Table 2.	Means of all dependent measures for problems in the transfer session of Experiment 1	55
Table 3.	Means of decision time for group size x problems in the transfer session of Experiment 1	56
Table 4.	Means of all dependent measures for problems in the learning session of Experiment 2	72
Table 5.	Mean decision times for each array in the learning session of Experiment 2	73
Table 6.	Mean decision times for arrays x problems in the learning session of Experiment 2	74
Table 7.	Mean decision times for each array in the transfer session of Experiment 2	75
Table 8.	Means of all dependent measures for problems in the learning session of Experiment 3	85
Table 9.	Means of all dependent measures for arrays in the learning session of Experiment 3	85
Table 10.	Mean trials to solution for arrays x problems in the learning session of Experiment 3	86
Table 11.	Mean untenable hypotheses for group size x vocalization in the learning session of Experiment 3	87
Table 12.	Mean strategy efficiencies for group size x vocalization in learning session of Experiment 3	88
Table 13.	Means of all dependent measures for problems in transfer session of Experiment 3	90

Table 14.	Means of all dependent measures for arrays in the transfer session of Experiment 3	90
Table 15.	Mean trials to solution for vocalization x problems in the transfer session of Experiment 3	91
Table 16.	Mean trials to solution for arrays x problems in the tranfer session of Experiment 3	92
Table 17.	Mean untenable hypotheses for group size x arrays in the transfer session of Experiment 3	93
Table 18.	Mean strategy efficiencies for arrays x problems in the transfer session of Experiment 3	94
Table 19.	Mean trials to solution for vocalization x grouped arrays	95
Table 20.	Mean trials to solution for grouped arrays x sessions	96
Table 21.	Mean untenable hypotheses for grouped arrays x sessions	98
Table 22.	Mean strategy efficiencies for grouped arrays x sessions	100
Table 23.	Mean decision times for grouped arrays x sessions	107
Table B1.	ANOVA table for trials to solution during learning Experiment 1	129
Table B2.	ANOVA table for untenable hypotheses during learning in Experiment 1	129
Table B3.	ANOVA table for strategy efficiency during learning in Experiment 1	130
Table B4.	ANOVA table for decision time during learning in Experiment 1	130
Table B5.	ANOVA table for trials to solution during transfer in Experiment 1	131
Table B6.	ANOVA table for untenable hypotheses during transfer in Experiment 1	131

Table B7. ANOVA table for strategy efficiency during transfer in Experiment 1 132

Table B8. ANOVA table for decision time during transfer in Experiment 1 132

Table B9. ANOVA table for trials to solution between sessions in Experiment 1 133

Table B10. ANOVA table for untenable hypotheses between sessions in Experiment 1 133

Table B11. ANOVA table for strategy efficiency between sessions in Experiment 1 134

Table B12. ANOVA table for decision time between sessions in Experiment 1 134

Table B13. ANOVA table for trials to solution during learning in Experiment 2 135

Table B14. ANOVA table for untenable hypotheses during learning in Experiment 2 135

Table B15. ANOVA table for strategy efficiency during learning in Experiment 2 136

Table B16. ANOVA table for decision time during learning in Experiment 2 136

Table B17. ANOVA table for trials to solution during transfer in Experiment 2 137

Table B18. ANOVA table for untenable hypotheses during transfer in Experiment 2 137

Table B19. ANOVA table for strategy efficiency during transfer in Experiment 2 138

Table B20. ANOVA table for decision time during transfer in Experiment 2 138

Table B21. ANOVA table for trials to solution between sessions in Experiment 2 139

Table B22. ANOVA table for untenable hypotheses between sessions in Experiment 2 139

Table B23.	ANOVA table for strategy efficiency between sessions in Experiment 2	140
Table B24.	ANOVA table for decision time between sessions in Experiment 2	140
Table B25.	ANOVA table for trials to solution during learning in Experiment 3	141
Table B26.	ANOVA table for untenable hypotheses during learning in Experiment 3	142
Table B27.	ANOVA table for strategy efficiency during learning in Experiment 3	143
Table B28.	ANOVA table for decision time during learning in Experiment 3	144
Table B29.	ANOVA table for trials to solution during transfer in Experiment 3	145
Table B30.	ANOVA table for untenable hypotheses during transfer in Experiment 3	146
Table B31.	ANOVA table for strategy efficiency during transfer in Experiment 3	147
Table B32.	ANOVA table for decision times during transfer in Experiment 3	148
Table B33.	ANOVA table for trials to solution between sessions in Experiment 3	149
Table B34.	ANOVA table for untenable hypotheses between sessions in Experiment 3	150
Table B35.	ANOVA table for strategy efficiency between sessions in Experiment 3	151
Table B36.	ANOVA table for decision time between sessions in Experiment 3	152
Table C1.	Mean decision times for groups size x arrays x problems in the learning session of Experiment 3	153

Table C2. Mean decision times for group size x vocalization x arrays x problems in the learning session of Experiment 3 154

Table C3. Mean decision times for group size x grouped arrays x session in Experiment 3 155

COOPERATIVE RULE-LEARNING: A METACOGNITIVE INTERPRETATION

The study of problem-solving in humans is of great importance. Understanding the processes of human problem-solving would be a boon to any institution in which people are faced with novel situations for which they must produce some unique solutions. In problem-solving, the person actively acquires and evaluates information which many be pertinent to the task. The relationships between items of information are restructured in various ways, with the intent of reaching a specific goal, the solution of the problem.

One popular laboratory analogue to problem-solving has been the rule-learning task. Rule-learning is an inductive problem-solving task. The learner evaluates available information concerning specific rules. This information generally takes the form of examples and nonexamples of a particular concept or rule. This information is restructured in various ways by the learner until a structure or rule is found which accounts for all of the data.

One attempt to understand the processing involved in such complex cognitive behavior is found in theories of metacognition (Flavell, 1976, 1978, 1979). Metacognition is the knowledge, experience, goals, and strategies individuals have which control their own cognitive functioning. These control processes help to determine the cognitive strategies individuals use in problem-solving. The general categories of person, task, and strategy variables which affect problem-solving are assumed to affect metacognition. These variables interact with metacognitive processes to determine individual problem-solving performance.

Other aspects of rule-learning which metacognitive theories may help to explain include the transfer of performance and the effects of learning with one or more partners. Through specific learning events, individuals may learn specific process characteristics about their own cognition, that of others, and that required by specific tasks. To the extent the characteristics of some initial task are similar to the characteristics of a transfer task, the same metacognitive control processes would be expected to be used.

When learning a task with other learners, those experiences which help to form the metacognitive control processes include influences from the presence and performance of the other learners. Thus, subsequent individual performance will be controlled by metacognitive control processes which were influenced by the presence and performance of others.

Effects of learning with others may be positive or negative, depending upon how the learning experience influenced the development of the metacognitive controls. If positive, learners should have improved cognitive monitoring of individual problem-solving. One method of requiring subjects to monitor their cognitive processing may be to require the subjects to vocalize their strategies during learning. Vocalization may require subjects to attend to the effectiveness of the cognitive strategies and the meaning of the informative feedback obtained during learning. In other words, required vocalization may be an observable performance analogue of the metacognitive process of monitoring the performance, feedback, and strategies. Subjects who are required to vocalize during learning may develop more positive metacognition than subjects who are not forced to monitor. Therefore, any positive effects of discussion during problem-solving may be due to the discussion functioning similarly to required vocalization in improving subsequent metacognitive monitoring.

The general purpose of conducting this series of studies was to investigate how group problem-solving affected subsequent individual problem-solving. These studies also included an investigation of whether or not vocalization of reasoning affected problem-solving and its transfer. Finally, it was discussed whether or not a metacognitive approach can account for the performance and transfer results.

RULE-LEARNING AS PROBLEM-SOLVING

When an individual faces a problem, the individual confronts a situation which cannot be solved by the simple application of any immediately available responses. Problems arise and must be solved in every aspect of human endeavor. Learning about the processes of problem-solving is important to education, industry, government, and business.

Any problem has three main characteristics. A problem consists of some given information, goals, and a variety of means for attaining those goals. The given information may be relevant as well as irrelevant. The information may be usable in its present form or many require restructuring. Often, the goals are not well-defined and must be determined before the means may be considered. Most frequently, the heart of the problem lies in determining the various means and evaluating which ones may be adequate. Just as the given information may need to be restructured, the available means may require alteration into new means. Thus, any problem can be considered to be composed of many smaller problems, the solution of which helps to determine the solution of the overall problem.

The process of problem-solving follows four major steps. The first step involves preparation. During preparation, the learner determines the critical characteristics of the givens and goals. The learner also prepares to attend to problem-solving skills. During reasoning, the second step, the learner actively and systematically restructures the problem. Relationships within the problem are redefined, and alternative solution strategies are reviewed, until a hypothesized solution is produced. The third major step involves the testing of this hypothesis. Feedback from the application of the hypothesis is used in evaluation, the last step. If the hypothesis satisfies the goals of the problem, the problem is solved. If, however, feedback indicates the problem is not solved, this information is added to the given and the process starts anew.

Rule-learning provides a good laboratory analogue of problem-solving. Haygood & Bourne (1965) identified three main categories of conceptual learning, called attribute identification, rule-learning, and complete learning. Concepts are abstractions used to categorize classes of objects or events. A specific concept is composed of various instances of events which are similar in their specific dimensions and attributes. These attributes are combined by specific rules which form the concept. In attribute learning, the subject is informed of the specific rule involved, as well as the number of attributes combined by the rule. The task is to correctly identify the specific attributes from an array of attributes presented simultaneously or successively. In rule-learning, the subject is told precisely which attributes are important or relevant. The task involves discovering the correct rule for combining these attributes. Finally, complete learning involves the discovery of both the specific attributes and rule for combining a known number of attributes. In this paper, rule-learning will be used to refer to both complete learning and rule-learning, as both involve the identification of the type of rule.

In each type of conceptual learning, the three major characteristics of problems are present. Subjects are given either an array of positive and negative instances, the number of relevant attributes, the rule, and/or the specific attributes. Their goal may be to identify the relevant attributes, general rule, or the rule and attributes. Means by which subjects attain their solutions are the foci of studies in this area.

Generally, the form the means take is also somewhat structured. Three paradigms have been used to structure these means, the simultaneous, selection, and reception paradigms. In a simultaneous paradigm, subjects are presented with the total array of instances, with the positive and negative instances labeled. Subjects make a series of hypotheses about the identification of the rule or attributes and receive confirming or disconfirming feedback for each hypothesis. In the reception paradigm, subjects are presented with a predetermined series of instances. For each instance, subjects predict whether it is positive or negative, and receive feedback on these predictions. In the selection paradigm, subjects are presented with the total array of positive and negative instances, without labels. Starting from a provided focus instance, subjects make hypotheses concerning the identification of the rule and/or select other instances for more feedback. Feedback indicates whether the selection was positive or negative and/or whether the hypothesis is correct or incorrect.

In sum, most problem-solving situations involve three characteristics: given information, goals, and some means for attaing the goals. Four major steps in problem-solving include preparation, reasoning, testing, and evaluation. Finally, rule-learning has often been used as a laboratory analogue to problem-solving. Within rule-learning, the means for attaining the goals of the problem have been structured by the use of simultaneous, reception, and selection paradigms.

Rule-Learning Perspectives

Historically, there have been two major theoretical perspectives for concept formation, a discrimination learning approach and a mediating processes approach (Kendler, 1961). Discrimination learning approaches assume the learner discriminates some specific aspects of the physical stimuli from the rest of the stimuli. The discriminated aspects form the concept. Mediational approaches assume some internal process mediates between the stimulus and the response. The most current mediational approaches are information-processing theories (Newell, Shaw, & Simon, 1958). From a general information-processing approach, the structure of the brain is postulated as functioning as if it were composed of many components, each responsible for certain aspects of the processing of information. General components include a sensory register, working memory, and long-term memory, containing episodic and semantic information. The long-term memory also contains various idiosyncratic strategies for problem-solving. The strategies are idiosyncratic in that the content of each strategy is determined by each individual's learning experiences. To the extent that learning experiences share certain features across individuals and contexts, the strategies are less idiosyncratic and may be categorized and compared. Strategies are influenced by both episodic and semantic memory. Information entering the long-term memory helps to determine which available strategies are evoked for use in a particular situation.

Two major sources of information are considered when a strategy is in use. New information from the stimulus and informative feedback is related to previously stored information. In rule-learning, this stored information consists of strategies and representations of possible solutions. Traditionally, these representations have been treated as if they consisted of defining attributes of concepts and rules (Bourne, 1966). New information is tested for relevant and irrelevant attributes and if a match between the stimulus characteristics and defining attributes of some rule is found, the problem is solved. More recently, it has been suggested that it is not the defining characteristics of concepts and rules which are stored, but prototypes of the concepts and rules (Glass, Holyoak, & Santa, 1979). Stimuli are then judged according to their similarity to the prototype. Whichever representation is correct is not important for the purpose of conducting these studies. Processing of new information, whether it be compared to some prototype or against some defining characteristics, can still be represented by hypothesis-testing.

Hypothesis-testing is the process by which an individual comes to make predictions concerning the categorization of new information. Specific hypothesis-testing strategies may vary from individual to individual, but overall processes of hypothesis-testing are assumed to be similar across individuals. The learner actively compares characteristics of the stimuli to a set of defining attributes or prototypes selected from long-term memory. On the basis of these comparisons, predictions or hypotheses are instantiated or altered.

Bruner, Goodnow, & Austin (1956) identified several strategies for categorizing the processing of subjects in various concept and rule-learning situations. Their four most famous strategies are those used in the selection paradigm with conjunctive rules. In simultaneous scanning, the learner tracks all potential solutions and chooses predictions to maximize the elimination of hypotheses. Depending on the complexity of the task, this strategy could cause a large memory load. The memory load is eased in successive scanning, as the learner sequentially tests hypotheses while tracking only the tested hypotheses. However, with this strategy, there is the danger of the learner testing from an incomplete or incorrect domain of hypotheses, thus increasing the memory load and trials to solution. Conservative focusing has been considered the most efficient strategy. In this strategy, the learner tests dimensions and attributes rather than full hypotheses. If a positive instance and a negative instance differ in just one dimension, the positive attribute of that dimension is critical to the solution. In conservative focusing the learner starts from a positive instance and changes one attribute in one dimension at a time. In the final strategy, focus gambling, the learner focuses as in conservative focusing, but changes the attributes of two dimensions at a time, in hopes of obtaining more information.

Johnson (1978) criticized Bruner's strategy identification techniques. Some of these criticisms could apply to any attempt at strategy identification. One problem with assigning individuals to specific strategy categories is that researchers assume that motivation of the learner is reflected by the strategy category. For example, classifying learners as focus gamblers assumes they are actually paying attention to all the altered dimensions and are motivated to use a more risky strategy in order to attain more information. Second, there is some evidence that learners are not consistent in their use of strategies. Finally, the use of strategy identification techniques is oriented more toward describing styles of processing, rather than in investigating the efficiency of the processing.

Laughlin (1978) used two measures of processing. The measure of strategy efficiency was called the focusing index, while the measure of the efficiency of the processing of feedback was called the proportion of untenable hypotheses. Johnson (1978, 1980) also developed a measure of processing efficiency called monitoring. Monitoring is measured by two variables, the number of instances testing previously known information and the number of untenable hypotheses. Measures of rule-learning processing may be related to the monitoring component of metacognition (Flavell, 1978, 1979).

In summary, the perspective used in this series of studies is a general information-processing approach. Learners are assumed to engage in a hypothesis-testing, in which they actively process previously stored and recently attended information in making predictions concerning the categorization of the new information. Researchers have attempted to identify different rule-learning strategies and to measure specific processes. A more in-depth discussion of how hypothesis-testing may reflect the information-processing of individuals in rule-learning is the topic of the following section. This is followed by a discussion of possible metacognitive control of hypothesis-testing.

Hypothesis-Testing and Rule-Learning

Hypothesis-testing was briefly defined as the process by which an individual comes to make predictions concerning the categorization of new information. The learner is cognitively active during hypothesis-testing, that is, the units of behavior are covert cognitive processes. Tumblin & Gholson (1981) distinguished the making of predictions from the use of response sets. The formation of predictions is responsive to feedback and tends to increase with age. Response sets are not responsive to informative feedback, tend to characterize the behavior of nonhuman animals and young children, and disappear with age (Levine, 1975; Tumblin & Gholson, 1981). Levine (1975) has identified three levels of hypothesis-testing theory, and assumptions concerning the learner at each level. These levels concern the hypothesis universe, the hypothesis domain, and the component processes of hypothesis-testing.

The universe of possible solutions consists of subgroups of domains of similar hypotheses. Conceivably, one hypothesis could be in many domains. Hypothesis-testing theory includes the assumptions that the learner recognizes the nature of the problem and can identify the domain of possible solutions. Brown (1975, 1978) indicated this metacognitive ability to recognize the nature of the problem improves with age. Depending on whether or not the learner fully understands the nature of the problem, the learner will sample from a correct or incorrect domain. A correct domain contains the solution hypothesis and all logical hypotheses, given the nature of the problem. Incorrect domains contain illogical or untenable hypotheses. A domain which contains logical hypotheses, but not the solution hypothesis, is an incomplete domain. By definition, a complete domain is also a correct domain.

Within each problem, a learner who selects from an incorrect domain learns nothing (Levine, 1975). All instances inferred from an incorrect domain will be negative. This does not mean that nothing can be learned from negative instances. Nothing is learned from any incorrect domain, except when the total domain has been tested and the learner discovers the domain was incorrect. The elimination of all hypotheses from one domain serves as a signal for the selection of a new domain.

Transfer was introduced by Levine (1975) in a discussion of between-problems dynamics. "When S receives a series of problems, he infers from the first N solutions the domain within the universe from which the (N+1)th solution will be taken. He will start the (N+1)th problem by sampling Hs from this domain" (Levine, 1975, p. 282). This transfer hypothesis may be extended to transfer other than from problem to problem within a single session. Such an extension might state that to the extent two problem-solving situations have similar characteristics, the solution hypotheses and strategies from the first situation will be applied in the second.

This more general transfer hypothesis can be applied to the various types of transfer discussed by Royer (1979). Vertical transfer occurs when the learner recognizes that a particular skill will contribute directly to the learning of a higher order skill. For example, learning to discriminate the number of sides on a geometric form contributes to learning the difference between a quadrilateral and a pentagon. Lateral transfer occurs when the acquisition of one skill helps in the acquisition of another, at the same level of complexity. Specific transfer occurs when there is a clear similarity between elements in the learning and transfer tasks. For example, learning to conjugate the French verb "finir" aids in learning the conjugations of the other -ir verbs. Non-specific or general transfer occurs when there are no obviously shared stimulus characteristics between the original and transfer tasks. This may occur if practicing solving various problems leads to a general problem-solving orientation when facing subsequent tasks. Near transfer occurs when the stimulus complex of the original learning and transfer tasks are similar. For example, the adding of two two-digit and two four-digit numbers involves the addition of all the numbers. In far transfer the stimulus complexes are less similar. For example, the learner may realize that adding or subtracting two two-digit numbers and balancing a checkbook both involve addition and subtraction.

At the level of the domain, the learner works with one subset of hypotheses at a time. From this subset, the learner chooses a hypothesis as the working hypothesis. Starting with the assumption that the working hypothesis is correct, the learner tests this hypothesis by selecting stimulus instances which alter one or more dimension of the working hypothesis. The learner receives feedback concerning whether or not the selected instance is in the set described by the rule to be learned. Feedback on each instance is also feedback about the appropriateness of the working hypothesis. Without feedback, the working hypothesis is maintained (Levine, 1975). If feedback confirms the plausibility of the working hypothesis, that hypothesis is repeated. If the learner disconfirms the working hypothesis, Levine (1975) assumes that the hypothesis is dropped and never tested again. However, it may be a bit strong to state that disconfirmed hypotheses are never tested again, as humans are not perfect information-processers. Disconfirmation of one hypothesis may be used to eliminate other hypotheses, if the learner also tracks these as well.

At the level of component processes, Levine (1975) suggests that the learner must first understand the nature of the problem well enough for the correct domain to be sampled. The learner selects the working hypothesis. This selection is determined by both task and learner characteristics. From the hypothesis, a stimulus instance is tested. The learner verbally encodes the chosen stimulus and rehearses all related hypotheses within the domain (Levine, 1975). Information from the trial and feedback session is encoded. If the selected instance is positive, the rehearsed hypotheses are encoded as possible solutions. These form the new subset of the hypothesis domain, and a new instance is tested. If the selected instance is negative, the rehearsed hypotheses are disconfirmed; and some hypothesis complementary to the initial working hypothesis becomes the new working hypothesis (Levine, 1975).

The retention of a working hypothesis upon confirmation has been called a "win-stay" strategy (Tumblin & Gholson, 1981). Tumblin & Gholson reviewed research which suggested that the "win-stay" strategy develops differentially during childhood. Early in childhood, response sets dominate. As the child develops, "win-stay" hypothesis-testing appears more frequently. However, not all adults use a totally "win-stay" processing procedure.

Developmental trends also appeared in hypothesis set size and sampling systems (Tumblin & Gholson, 1981). As the child develops, the number of hypotheses the child tracks at one time increases. Young children use zero-memory and local consistency memory strategies. In zero-memory hypothesis-testing, the learner tests one hypothesis at a time, returning disconfirmed hypotheses to the sampling pool. Working memory is required to track only one hypothesis at a time. In local consistency hypothesis-testing, the learner tests only hypotheses from an array of hypotheses based on all positive instances. Thus, a negative instance results in the disconfirmation of any hypothesis which could have indicated that choice. This memory strategy requires the learner to retain all hypotheses indicated by positive instances and remember which instances have been tested. Elementary school children tend to use zero-memory and some local consistency hypothesis-testing, while adults use more local consistency and some global consistency hypothesis-testing (Tumblin & Gholson, 1981). Global consistency hypothesis-testing assumes the learner uses all the information from both positive and negative instances to alter the hypothesis domain. This memory strategy can be considered a perfect processing model. Implicit in this strategy is the memory tracking of multiple hypotheses and instances. Finally, there is a developmental trend in sampling techniques. Kindergarten children tend to use stereotyped sampling techniques which resemble response sets, while children in later elementary grades use strategies which test each attribute and/or dimension separately. College students tend to use more focusing strategies, indicating the use of global consistency (Tumblin & Gholson, 1981).

Levine (1975) discussed three levels of hypothesis-testing: the levels of the hypothesis universe, domain, and components. At the level of the hypothesis universe, successful and efficient hypothesis-testing requires learners to recognize the nature of the problem and to identify the complete domain of possible solutions. Within each hypothesis domain, learners use feedback from tested hypotheses to confirm or disconfirm possible solutions. Hypothesis-testing in solving specific problems occurs at the components level. Transfer was discussed as the continuation of specific hypothesis-testing strategies from problem to problem (Levine, 1975) and context to context (Royer, 1979). Developmental trends in hypothesis-testing were also discussed.

Metacognition

One developmental theory of control processes is the theory of metacognition discussed by Brown (1975, 1978) and Flavell (1976, 1978, 1979). Metacognition is all the knowledge and beliefs an individual has come to hold concerning cognitive processing and personal cognitive functioning (Flavell, 1976). As developmental changes occur at the level of cognitive processing, changes also occur in metacognitive processing. That is, there are developmental changes in the control processes of problem-solving and hypothesis-testing. Brown (1978) and Flavell (1979) briefly reviewed evidence of the developmental nature of metacognition. Young children were usually not aware of when they had studied long enough to memorize a set of items. Children were also not aware of inadequacies in obscure verbal instructions. "Results such as these have suggested that young children are quite limited in their knowledge and cognition about cognitive phenomena, or in their metacognition, comprehension and other cognitive enterprises" (Flavell, 1979, p. 906). Metacognition develops through an interaction of physiological development and learning experiences. Thus, a young child may not be aware of the inadequacies of the teacher's instructions because the child's appropriate physiological structure has not developed and/or because the child has not had the experience of monitoring the teacher's instructions for obscurities and omissions. Given physiological readiness and experience at handling inadequate instructions, the child develops the metacognitive knowledge that for a task to be completed, the instructions must be clear and complete.

Model of cognitive monitoring

Metacognitive knowledge is one component of Flavell's model of cognitive monitoring (Flavell, 1979). The full model consists of an interaction of metacognitive knowledge, metacognitive experiences, task goals, and strategies. Cognitive monitoring involves both metamemory and a sensitivity to the existence of a problem (Flavell, 1978). Metamemory contains all the knowledge and beliefs individuals have concerning their storage and retrieval capabilities. Sensitivity includes an awareness of which problems do or do not require intentional memory work for a solution.

Metacognitive knowledge is the knowledge and beliefs, held by an individual in long-term memory, about what factors affect that individual's cognitive efforts. This knowledge may be activated intentionally or automatically. Factors involved in metacognitive knowledge include person, task, and strategy information. Person variables include information concerning the individual's strengths and weaknesses in various tasks; as well as their strengths and weaknesses in relation to those of others. Person variables also contain information which pertains to problem-solving in general. These are called universals of cognition by Flavell (1979). Finally, person variables include metacognitive monitoring processes. These are processes involved in interpreting the progress of ongoing cognitive monitoring. In other words, metacognitive monitoring is the executive control of cognitive monitoring.

The remaining factors in metacognitive knowledge are task and strategy variables. Task variables include evaluation of the quality and quantity of the information available in a task. The learner knows what task factors influence the storage and retrieval of information. Other task information includes the mnemonic value of specific relationships and strategies. Strategy variables are involved in evaluating the effectiveness of various strategies in respect to the goals and context of the task.

The second component of Flavell's model of cognitive monitoring is metacognitive experiences. Metacognitive experiences involve how the learner feels about ongoing cognitive activity. Individuals may feel positive or negative about their current processing. These experiences vary in duration and complexity. An individual should be most aware of metacognitive experiences during tasks which require intensive cognitive processing. Metacognitive experiences may lead to continued use or termination of cognitive strategies.

The last factors of cognitive monitoring are the goals of the task and the strategies employed in attaining those goals. Goals help determine what metacognitive knowledge is relevant and what strategies may be required. The progress of these strategies influences the metacognitive experience which, in turn, influences the revision or further use of the cognitive strategy.

Cognitive strategies are invoked to make progress in problem-solving; metacognitive strategies monitor this progress (Flavell, 1979). This interaction is important for understanding how cognition and metacognition influence each other. For example, if a learner has matured to a stage appropriate for the development of some form of metacognitive monitoring, the problem-solving experiences this learner has will help form the metacognitive control of future problem-solving. This interaction of experience and metacognitive monitoring does not stop in adulthood. An adult learner is capable of improving metacognitive control and future cognitive performance given the appropriate experiences.

Improving metacognitive control

Brown (1975) described a model for studying the development of memory. This model can be adapted to describe a model for the study of aiding adult learners in improving their metacognitive control and cognitive strategies. When working on problems in which "mnemonic strategies are required, the developmentally young should perform poorly compared with more mature subjects, for they fail to employ mnemonic strategies effectively" (Brown, 1975, p. 137). Rule-learning is a strategic semantic task within this model. The first step in measuring learner performance in such tasks is to measure the appropriateness of any spontaneously adopted strategy. This amounts to making inferences regarding the adequacy of the metacognitive controls already in existence. In rule-learning tasks, this is inferred from measures of the processing of feedback, strategy efficiency, and the number of trials to solution. If the learner's spontaneous strategies are not adequate or could be improved, the next step is to attempt to induce the use of a better strategy. The metacognitive analogue to this step is to provide the learner with adequate experiences to induce better metacognitive control and processing. Perhaps the reason problem-solving with a partner has often aided rule-learning performance (Johnson, Maruyama, Johnson, Nelson, & Skon, 1981) is that some aspect of cooperative problem-solving can provide the adequate experience for improved metacognitive control.

If the learner cannot be induced to use a better strategy when one is logically available, the learner may have a mediational deficiency (Brown, 1975). However, if the learner uses a better strategy as a result of some experimental prompting, the learner may have had a production deficiency (Brown, 1975). The strategy was available to the learner, but the metacognitive control was not adequate to elicit the appropriate strategy.

Assuming the learner either spontaneously adopted an appropriate strategy or had a production deficiency, the model can be extended to test whether or not the learner can solve subsequent problems with equal efficiency. Metacognitive theory would predict that if experimentally induced experiences were adequate to promote effective metacognitive control in subjects with production deficiencies, the performance of these subjects should be nearly the same as that of subjects who spontaneously adopted the appropriate strategy.

Vocalization of reasoning

One experience which may fit this model and facilitate cognitive monitoring is learner vocalization of reasoning during problem-solving (Davis et al., 1968; Durling & Schick, 1976; Eifermann, 1965; Gagne & Smith, 1962). Some researchers have suggested that discussion in cooperative problem-solving groups can be an important variable for improving performance (McGlynn & Schick, 1973a; Schick & McGlynn, 1976). Durling & Schick (1976) operationalized cooperative discussion as vocalizing one's reasoning to a peer. They found vocalizing pairs used fewer trials to solution, were superior in evaluating feedback, and used more efficient testing strategies than nonvocalizing pairs.

The use of subject vocalization of their own reasoning to aid performance can be traced to Gagne & Smith (1962). They found ninth and tenth grade boys who were required to vocalize on a 3-ring Tower of Hanoi problem performed better on larger sets than nonvocalizing boys. Also, although vocalizing required more time, subjects who had vocalized were faster on a subsequent test which required no vocalization. Using a similar task with college students, Davis et al. (1968) found similar results for a final nonvocalized task. These researchers also found the presence of the experimenter facilitated performance, especially during practice vocalization tasks.

Eifermann (1965) studied the vocalization of first-year psychology students on a concept attainment task completed via a selection paradigm. Eifermann discovered subject justification of their instance selections or tests was positively related to efficient concept attainment. Using a similar task, Laughlin & Doherty (1967) found pairs of college females allowed discussion, used fewer trials to solution, repeated hypotheses, and illogical hypotheses than pairs not allowed to discuss. However, discussing pairs required more time to solution. McGlynn & Schick (1973a) found cooperative discussing pairs of college students used fewer trials to solution, more efficient strategies, and better evaluation of feedback than cooperative nondiscussing pairs and individuals, on a similar rule-learning task.

McGlynn (1972) used a similar task, but with groups of four cooperating subjects, competing pairs, and individuals. Cooperative quads used fewer trials to solution, fewer illogical hypotheses, and more focusing than individuals. Competing pairs performed similarly to individuals for the proportion of untenable hypotheses, and similarly to cooperative quads for focusing. In other words, cooperative pairs who competed with other pairs did not evaluate feedback any better than individuals, but did develop an efficient testing strategy. Schick & McGlynn (1976) suggested that cooperative pairs competing with other pairs performed as well as cooperative pairs cooperating with other pairs if all groups are allowed the same amount of discussion. Thus, discussion may be a more salient variable than cooperation versus competition in affecting college level rule-learning.

Durling and Schick (1976) asked vocalizing pairs to thoroughly discuss their card choices and hypotheses in a concept attainment task. Performance of these vocalizing pairs was compared with that of individuals vocalizing to the experimenter or a peer, nonvocalizing pairs, and nonvocalizing individuals. Vocalizing pairs and individuals vocalizing to a peer used fewer trials to solution than nonvocalizing pairs. Vocalizing pairs also used more efficient strategies than nonvocalizing pairs. Vocalizing to an experimenter was found to be no more efficient than working alone without vocalizing. Vocalizing pairs were the best at processing feedback. Vocalizing individuals were better processers of feedback than nonvocalizing pairs. These results suggest that vocalization of reasoning may be one characteristics of working in small groups which accounts for small group superiority.

Implicit in the Durling & Schick study and the above suggestion is the assumption that vocalization in groups includes verbal expression of individual reasoning. Real life groups may have discussion devoid of any mention of individual reasoning. If a group is told they may discuss, but not explicitly told to discuss their reasoning, discussion may or may not include verbalization of the reasoning. This series of studies accounts for this possibility by comparing the performance of individuals and cooperative pairs allowed to discuss, with the performance of individuals and cooperative pairs who are allowed to discuss and asked to explain their reasoning.

Within the framework of metacognitive theory, verbalization of reasoning would aid in the problem-solving process in two ways. First, performance during verbalization tasks may improve, as verbalization of reasoning may act as an active cognitive monitoring system. Second, verbalization of reasoning should improve the metacognitive monitoring of future problem-solving endeavors. These studies also provided a means of investigating this effect. Two days following the initial rule-learning task, all subjects individually completed a similar rule-learning task, without being required to verbalize. If verbalization of reasoning did improve metacognitive monitoring, the problem-solving of subjects who initially verbalized should be superior to that of subjects who were not required to state their reasoning in the initial task.

Summary

Metacognition is the beliefs and knowledge learners have concerning their cognitive functioning (Flavell, 1976). Flavell also developed a model of cognitive monitoring, consisting of metacognitive knowledge, metacognitive experiences, task goals, and strategies. Cognitive monitoring controls the progress of problem-solving. Metacognition controls the cognitive monitoring. It may be possible to improve individual problem-solving by improving cognitive monitoring. Subjects' vocalization of reasoning during problem-solving was suggested as a possible means of improving the cognitive monitoring.

FACTORS WHICH INFLUENCE RULE-LEARNING

This section constitutes a brief review of those factors which have been demonstrated to influence rule-learning. These factors can be subsumed under two headings, task and person variables. Task variables include characteristics of the context, physical stimuli, memory requirements, and feedback. Person variables include developmental level, personality characteristics, and sex.

Task Variables

Task variables include those characteristics of any rule-learning task which may affect the performance of learners in that task. Context characteristics include method-related and person-related variables. Method-related variables are those variables which differ with various procedures. Person-related variables also depend upon the procedure, but are more directly related to the subjects than are method-related variables. A second major class of task variables are stimulus characteristics. These include physical characteristics of the array and those of the stimulus selections. Required memory characteristics form the third class of task variables. These determine the amount of memory work required of the learner. The final class of task variables is feedback characteristics. The timing and quality of the feedback provided for the learner help to determine the rule-learning performance.

Context characteristics

The paradigm, type of rule, and amount of pretraining are all method-related variables. Person-related variables include group size, cooperation versus competition, and grouping by ability.

Paradigms Three major paradigms have been described, the simultaneous, selection, and reception paradigms. The selection and reception paradigms have been the most widely used. Researchers interested in hypothesis-testing have tended to use the selection paradigm. Results of studies comparing these two paradigms indicate selection paradigms provide for the best performance with older subjects and less complex rules (Laughlin, 1969). Laughlin compared the problem-to-problem transfer for these paradigms and found interproblem transfer was related to the difficulty of the problem, more than to the type of paradigm. However, the subjects given the reception paradigm required fewer trials to solution for more complex problems and used a lower proportion of untenable hypotheses than subjects given the selection paradigm (Laughlin, 1969, 1972).

Type of rule Neisser & Weene (1962) distinguished between three levels of rules. These levels were found to be hierarchically arranged such that the more complex, higher level rules were more difficult to attain than the lower order rules (Braley, 1963; Conant & Trabasso, 1964; Haygood & Bourne, 1965; Laughlin, 1968, 1969; Laughlin & Jordan, 1967; Laughlin et al., 1968b; Neisser & Weene, 1962. Most rule comparisons were between biconditional relations, conjunctive, and disjunctive rules. Biconditional rules take the form of "A if and only if B" (Haygood &

Bourne, 1965). Conjunctive rules take the form of "A and B." Disjunctive rules have two forms, inclusive and exclusive. The former is "A or B, or both;" the latter, "A or B, but not both." Most disjunctive rules in research have been inclusive.

Conjunctive rules have been consistently easier to learn than disjunctive rules (Conant & Trabasso, 1964; Laughlin, 1968, 1969; Laughlin & Jordan, 1967; Laughlin et al., 1968b). However, biconditional rules have been found to be similar in difficulty to conjunctives (Laughlin, 1969; Laughlin & Jordan, 1967). Laughlin & Jordan found subjects used a similar number of trials to solution, time to solution, and scanning strategy for conjunctive and biconditional rules. More focusing strategy was used with conjunctive rules.

Pretraining Pretraining subjects on the type of rule to be later tested has been suggested as one method of facilitating near transfer (Kendler, 1961). Wells (1963) demonstrated support for this assertion by indicating that pretraining subjects on disjunctive rules led to more disjunctive solution attempts on subsequent problems. However, the effect of pretraining is not clear-cut, especially when ability groups of different sizes are pretrained. Lemke, Randle, & Robertshaw (1969) found an interaction of group size and ability at low levels of pretraining. Subjects were pretrained in groups of two or four, or as individuals. All subjects individually solved two rules following the pretraining. With a low level of pretraining, low ability subjects learned conjunctive rules faster after being homogeneously paired for ability. However, high ability subjects, with the same level of pretraining, learned the rules faster following pretraining in homogeneous groups of four, or as individuals. Lemke & Hecht (1971) used a similar procedure with individuals and pairs, and three levels of pretraining. As pretraining increased, the subsequent time to solution decreased. Results similar to Lemke et al. (1969) were also reported.

Other method-related context characteristics include the distribution of practice and the proximity of the experimenter. While investigating the role of memory in rule-learning, Dominowski (1965) found evidence favoring distributed practice. Davis et al. (1968) found the immediate presence of the experimenter facilitated learning the performance rule on a five-disk Tower of Hanoi task.

Group size Group size is the first person-related context characteristic. Although individuals do not logically constitute a group, research involving the use of individuals and groups will be discussed in this section. Effects of individual and small group problem-solving on rule-learning performance has been studied by many researchers. Most of this research can be subsumed under two categories. The first category concerns strictly performance studies in which the rule-learning performance of small groups and individuals is compared. The second is concerned with retention and transfer, in which the effects of small group and individual rule-learning on subsequent individual rule-learning is examined.

A recent meta-analysis of small group and individual problem-solving indicated that, in general, the problem-solving product of small groups is superior to that of individuals (Johnson et al., 1981). Results of rule-learning performance studies support this conclusion (Laughlin, 1965, 1972; Laughlin et al., 1968a; Laughlin et al., 1968b; Laughlin & Sweeney, 1977; McGlynn, 1972). Pairs of subjects have been found to use fewer trials to solution, more focusing, and a lower proportion of untenable hypotheses than individuals. McGlynn (1972) found cooperative groups of four were consistently superior to competing individuals on all of the above measures. However, cooperative pairs competing with other pairs used a proportion of untenable hypotheses similar to that of individuals.

Not all performance studies found clear main effects favoring small group performance (Hill, 1982). Groups required more time to solution than individuals (Laughlin, 1965). Schick & McGlynn (1976) balanced the opportunity to freely discuss in groups, and found no group size differences. They suggested that discussion may be a more powerful variable in problem-solving than is the cooperativeness of the group. Durling & Schick (1976) further investigated the effects of group size and discussion. They found pairs of subjects, allowed to state their rule-learning reasoning, outperformed pairs not allowed any discussion.

Results of transfer studies are not as clear as those of performance studies. Most of the transfer studies have used time to solution as the dependent measure. Klausmeier, Wiersma, & Harris (1963) found individually trained subjects were faster than individuals from pairs or quads on a transfer task of identifying one conjunctive rule, following a 30-second retention interval. However, there was no difference in performance between the groups following a 12-minute interval. Lemke et al. (1969) found subjects, initially trained in groups, individually were slower than individually trained subjects on a transfer task. Beane & Lemke (1971) found high ability subjects, grouped homogeneously by ability, performed faster after pretraining in pairs, while low ability subjects performed faster after pretraining in groups of four. They interpreted these results to suggest low ability subjects have less interference from idiosyncratic strategies, while adopting a group strategy, than do higher ability subjects. Finally, Lemke & Hecht (1971) found no significant results involving group size. However, low ability subjects tended to perform best following pretraining in pairs rather than individual training.

Dependent variables other than time to solution have been used in some rule-learning transfer studies. Laughlin & Sweeney (1977) found individual subjects used more trials to solution in the transfer task than did subjects who performed the transfer task in groups of three. Bender (1980) found individual and paired subjects tested individually, did not differ in time per trial in the transfer task. However, paired subjects used less time per trial in the transfer task than they did in the initial task. Individual subjects used more time per trial in the transfer task than in the initial task.

Cooperation McGlynn (1972), McGlynn & Schick (1973a), and Schick & McGlynn (1976) represent a series of studies investigating the impact of cooperation versus competition and the importance of discussion for rule-learning. As stated earlier, McGlynn (1972) found cooperative pairs competing with other pairs used efficient strategies, but were less skilled at processing and applying informative feedback than were cooperative groups of four. McGlynn & Schick (1973a) investigated the importance of discussion in cooperative pairs by comparing the performance of pairs allowed discussion, pairs not allowed discussion, and individuals. Pairs not allowed discussion performed similarly to individuals on all measures. Pairs allowed to discuss required fewer trials to solution, a lower proportion of untenable hypotheses, more focusing, and more time to solution. Finally, Schick & McGlynn (1976) studied the effects of private or public discussion in cooperative and competitive groups. Public discussion was discussion which could be overheard by others. They suggested that discussion is a more powerful determinant of group performance than the level of cooperation.

Ability groups One method of grouping subjects in the laboratory and classroom has been to group the individuals of similar ability together. Beane & Lemke (1971), Lemke & Hecht (1971), and Lemke et al. (1969) studied the effects of grouping by ability on rule-learning performance and individual transfer. Higher ability subjects generally outperformed lower ability subjects. Lemke et al. (1969) found low ability subjects in homogeneous pairs, briefly pretrained, performed best after training in pairs. However, the time to solution for low ability subjects trained in groups of four and for those trained as individuals was nearly the same following extended training. Beane & Lemke (1971) found no interactions involving both ability and the degree of pretraining. However, individuals from homogeneously grouped high ability pairs and low ability quads performed better than high ability quads or low ability pairs. This result is seemingly at odds with those of Lemke et al. (1969). Some difference in the composition of the ability groups in each study may account for the differing results. Finally, Lemke & Hecht (1971) found low ability subjects, with a low level of pretraining, tended to perform better as individuals following paired, rather than individual pretraining. High ability subjects under the same level of pretraining, tended to perform better following individual pretraining.

Summary Many characteristics of the rule-learning context affect the outcome. Selection paradigms may be the most appropriate for older subjects and less complex rules, such as conjunctives. Pretraining tends to improve rule-learning performance. However, levels of pretraining interact in a complex manner with group size and ability. Groups tend to outperform individuals on most measures of rule-learning. However, groups tend to be slower. Finally, group superiority may not transfer to subsequent individual performance. The initial advantage of group rule-learning appears to stem from a discussion of reasoning which may characterize the processing of cooperative groups.

Stimulus characteristics

Stimulus characteristics include the physical characteristics of the array and/or instances. Characteristics of the array include the dimensions in the array, its organization, and the sensory modality by which the array is perceived. Characteristics of the instances include the amount of redundant information and negative instances.

Array characteristics Array characteristics include the type of visual stimuli (Laughlin & Doherty, 1967), organization of the array (Laughlin, 1965), and reception of the array in various sensory modalities

(Laughlin et al., 1968a). Information from a complex stimulus array is difficult to encode, especially if the complexity does not add to the amount of relevant information (Bourne, 1966). However, if all of the attributes of all of the dimensions of the array were equally distinguishable, they may be expected to transmit equal amounts of information, and be equally easy to encode and classify (Archer, Bourne, & Brown, 1955). Archer et al. (1955) found the dimensions of size, form, and brightness to be equally distinguishable for use in rule-learning tasks.

Laughlin & Doherty (1967) found no effect for form versus sequence of symbols arrays for the number of trials to solution or repeated hypotheses. Subjects given a form array used fewer untenable hypotheses than subjects given the sequence array. Also, four-attribute problems required more time and elicited more untenable hypotheses than two-attribute problems. Laughlin (1965) found form arrays and four-attribute problems required more focusing strategies than sequence arrays and two-attribute problems, respectively. Laughlin also found no performance differences between ordered and random arrays.

Laughlin et al. (1968a) discussed evidence indicating no differences between visual, auditory, and mixed modes of presentation for single value rules. However, the mixed mode was less effective for bivalue rules. Visual arrays required more instances to solution than auditory arrays, which required more than mixed presentations.

<u>Redundant information</u> Effects of being presented redundant information and irrelevant information shall be discussed together. In order to control the amount of irrelevant information presented to each

learner, most studies have included the reception paradigm. Bourne (1966) suggested that redundant information may be positive to the extent it is also relevant to the solution of the task. Kendler (1961) reported the inclusion of nonredundant, irrelevant information negatively affected performance more so than the inclusion of redundant, irrelevant information. Thus an apparent hierarchy develops, with redundant relevant information being the least negative, followed by redundant irrelevant, and nonredundant irrelevant information.

Adding irrelevant information to the task negatively affects rule-learning performance in various ways. Archer et al. (1955) found additional irrelevant information increased the time subjects required to reach criterion. Processing errors before solutions also increased with additional irrelevant information (Archer et al., 1955; Hunt, 1961; Osler & Trautman, 1961). Osler & Trautman (1961) assumed that if hypothesis-testing were more frequent among the better subjects, added irrelevance should affect their performance more than that of the less able subjects. Results indicated that superior subjects performed on the same level as less able subjects, when irrelevance was added to the stimuli.

Negative instances   Subjects tend to have difficulty performing rule-learning tasks when using only negative instances (Bourne, 1966; Braley, 1963; Hovland & Weiss, 1953; Smoke, 1933). A negative instance is a selection from an array, which is not included in the set described by the rule to be learned. Researchers investigating the use of information from strictly negative instances have used all three paradigms. Studies of negative instances can be traced to Smoke (1933). Smoke used a simultaneous paradigm with only positive or both positive and negative instances labeled. Although there was no difference in the time required to learn from each array, most subjects preferred the array containing both positive and negative instances labeled. Hovland & Weiss (1953) extended Smoke's results by using the reception, simultaneous, and combined methods of presentation. Labels were on all positive instances, all negative, or both positive and negative instances. The dependent measure was the percentage of problems solved. The best performance occurred with all positive instances labeled, followed by the combined positive and negative instances, then the negative instances.

A second source of interest in negative instances stems from Bruner, Goodnow, & Austin (1956). In describing various selection strategies for conjunctive rules, Bruner, Goodnow, & Austin stated the subject would perform best by observing what changes in attributes caused a positive instance to become negative. For disjunctive rules, the subject focuses on a negative instance and observes what changes cause it to become positive. Both entail using information from negative instances. Disjunctive rules would require a greater use of negative instances than would conjunctive rules. Conant & Trabasso (1964) compared the learning of conjunctive and disjunctive rules on a selection paradigm. Conjunctive rules were easier, but with extended practice, subjects did learn to use the negative focus for disjunctive rules quite effectively. Freibergs & Tulving (1961) found extended practice with negative instances in a reception paradigm also improved subject performance.

Many reasons have been offered to explain why subjects have difficulty with negative information. Smoke (1933) suggested negative instances are of most value in preventing premature conclusions.

> It appears that in so far (sic) as negative instances assist concept learning they do so largely because of the way in which they prevent the learner from coming to one or more erroneous conclusions while he is still in the midst of the learning process (Smoke, 1933, p. 588).

This may be the case in simultaneous paradigms, but does not explain the full function of negative instances in other paradigms. In reception paradigms, negative instances may provide irrelevant information, as well as prevent premature conclusions. Negative instances provide the most relevant information in selection paradigms using conjunctive rules, and serve as foci for disjunctive rule-learning. Bruner, Goodnow, & Austin (1956) suggested subjects have been more exposed to positive instances in the environment. Thus the "inability" to use negative instances as well as positive instances may be due to lack of practice. Finally, Kendler (1961) suggested negative instances many cause more memory load and, therefore, be used less efficiently.

Summary Characteristics of the stimulus and instances may affect rule-learning. Highly complex rules are more difficult than less complex rules. However, the complexity of the stimulus array itself may not always affect rule-learning. Visual arrays may be more difficult to use than auditory arrays for complex rules. Added irrelevance to the stimulus instances negatively affects rule-learning, especially if it is not redundant. Finally, subjects have difficulty using strictly negative instances.

Memory characteristics

A number of memory manipulations have been found to affect rule-learning. Requiring the subject to memorize the array and to perform all rule-learning "in the head" led to poor strategies and redundant hypothesis-testing (Bourne, 1966). Memory manipulations which lessen possibly interference from irrelevant sources of information improved performance (Dominowski, 1965; Hunt, 1961).

One of the most frequent memory manipulations has been to allow some subjects to keep track of tested hypotheses and/or instances by use of paper and pencil or moving tested instances to separate piles. Bourne (1966) suggested the availability of previous information facilitated rule-learning. This was demonstrated by Cahill & Hovland (1960), who found subjects allowed unlimited access to information from previous selections correctly identified a greater proportion of rules than subjects who were only allowed to review the preceding selection. Moving each selection to an appropriate category upon feedback resulted in fewer trials to solution (Laughlin, 1969) and a lower proportion of untenable hypotheses (McGlynn & Schick, 1973a). Laughlin (1968) found no difference in the amount of focusing used by subjects in a paper and pencil versus no aid condition. However, Laughlin & Doherty (1967) found the paper and pencil memory aid resulted in fewer repeated hypotheses. Finally, females may make greater use of paper and pencil memory aids to reduce the proportion of untenable hypotheses than do males (Laughlin et al., 1968b).

Feedback characteristics

Feedback characteristics are related to memory requirements in that less informative feedback can increase memory load. Other feedback characteristics, besides the amount of information, are also related to memory load. Feedback which confirms irrelevant dimensions inhibits performance, as does infrequent feedback (Bourne, 1966). Dominowski (1965) found the post-response delay of feedback in rule-learning may not affect performance, furthermore, increased post-feedback delay may facilitate performance.

Other feedback characteristics include the type of instance, the form of the feedback, and probability of attaining correct feedback. Subjects' difficulty in dealing with strictly negative instances has been discussed. However, with practice, subjects can learn from negative instances. Tumblin & Gholson (1981) reported children learn better from verbal or symbolic feedback than from tangible feedback, such as tokens. The completeness of the informative feedback also affects rule-learning (Bourne, 1966). This was supported by Laughlin et al. (1968b) who reported superior performance with feedback describing instances as positive, negative, or partially positive, over feedback consisting of simply positive or negative labels. The rules used in this study were relational rules which allowed the use of "partially positive" as feedback. Probabilistic feedback of less than 100% correct led to increased guessing (Bourne, 1966). Finally, Tumblin & Gholson (1981) reported feedback on incorrect trials tended to be superior to feedback for correct responses only.

Feedback has been used as a design characteristic other than as a variable. McGlynn & Schick (1973b) and Johnson (1978, 1980) used a form of feedback designed to provide the subject with as little information as possible. This was accomplished by consistently labeling selections as either positive or negative, so as to eliminate as few hypotheses as possible on each trial. Thus, the solution is the hypothesis remaining after all other hypotheses have been eliminated.

Person Variables

Developmental trends have been reported in rule-learning and hypothesis-testing (Levine, 1975; Tumblin & Gholson, 1981), and memory processes (Brown, 1975, 1978; Flavell, 1978, 1979; Hagen, 1975). Flavell (1978, 1979) suggested a developmental component of metacognitive processing also exists.

The development of efficient rule-learning is age-related (Osler & Fivel, 1961; Osler & Trautman, 1961; Osler & Weiss, 1962). Osler & Fivel compared the rule-learning of 6, 10, and 14 year-old subjects using a stimulus-response approach. Errors decreased and the number of subjects who attained the concepts increased with increased age. Osler & Trautman (1961) suggested older subjects tended to use more hypothesis-testing than younger subjects.

Memory and metacognition also show developmental trends. Hagen (1975) reported subjects' spontaneous use of rehearsal strategies was age-related. Hagen also noted the accuracy of children's estimates of their memory span increased with age. Flavell (1978, 1979) reported that self-monitoring and the use of strategies develops with age. Five age-related components of the use of strategies include an awareness that retrieval is needed, where to focus problem-solving efforts, a systematic memory search, sequencing plans within a strategy, and using cues as an aid to retrieval (Flavell, 1978).

Tumblin & Gholson (1981) also reported developmental trends in the use of hypothesis-testing strategies. They compared this development to the cognitive-developmental stages of Piaget. Children in a preoperational stage tended to use stereotyped response patterns or sets rather than hypothesis-testing. Simple dimension checking and hypothesis-checking emerged during the concrete operations stage. Focusing occurred most frequently in the formal operations level (Tumblin & Gholson, 1981).

<u>General mental ability</u>

General mental ability and/or intelligence have been found to be positively related to rule-learning performance. Individuals with higher IQs tend to be more rapid learners than individuals with lower IQs (Bourne, 1966; Tumblin & Gholson, 1981). Osler & Fivel (1961) found hypothesis-testing in children was a function of IQ. More children with higher IQs reached the criterion of 10 consecutive correct responses than lower IQ children. Osler & Trautman (1961) supported the idea that higher IQ children were more frequently hypothesis-testing. Children with initially superior rule-learning performance lost all advantage when the array from which they were learning contained additional irrelevant attributes. Osler & Trautman suggested that these irrelevant attributes interfered with the hypothesis-testing of the higher IQ children. Osler & Weiss (1962) suggested that one source of the superiority of higher IQ children was their ability to supplement instructions. When specific problem-solving instructions were included, the effects of IQ disappeared, while that of age remained.

Many of the studies of general mental ability in college students have involved grouping the subjects on the basis of ability (Beane & Lemke, 1971; Lemke & Hecht, 1971; Lemke, Randle, & Robertshaw, 1969). In these studies, subjects were trained individually or in groups, but later tested as individuals. All results pertain to the individual performances. High ability subjects performed faster after training in mixed ability groups, while low ability subjects performed faster following training in homogeneous groups (Beane & Lemke, 1971). Following low levels of training, low ability subjects performed faster if their training had been in pairs, while high ability subjects tended to respond best to individual training (Lemke & Hecht, 1971; Lemke et al., 1969).

Personality variables

Few studies have investigated sex and personality differences in rule-learning. Bourne (1966) reviewed the inverted-U interaction of anxiety and task difficulty. Denny (1966) found intelligent, high anxiety subjects learned more rules than less intelligent, high anxiety subjects. Subjects with low anxiety did not differ between low and high intelligence levels. Tumblin & Gholson (1981) reported subjects with a reflective response style generally had longer response latencies and made fewer errors than less reflective subjects. Finally, McGlynn & Schick (1973a) found females used less time to solution than males.

PURPOSE

The general purpose of conducting these studies was to investigate the effects of rule-learning in pairs on subsequent individual rule-learning. Rule-learning was studied from a hypothesis-testing approach, in which the subject was seen as evaluating feedback about the task and applying this feedback to the development of a rule-learning strategy. Thus, transfer would be demonstrated in the development over trials of an effective, efficient strategy and in the application of the strategy to a new situation. Therefore, the focus of these studies was on the development and use of efficient strategies rather than specific strategy styles.

A more specific purpose of conducting these studies was to determine if support existed for the suggestion that effective discussion, defined as vocalization of the problem-solving reasoning, was a more powerful variable than simply working with others. If this assertion was correct, it was expected that subjects who reasoned aloud during training would have a more efficient strategy, be better at processing feedback, and use fewer trials to solution than subjects who did not verbalize. This would be true for both paired and individually trained subjects.

EXPERIMENT ONE

Experiment 1 is a reanalysis of the Bender (1980) study of the effects of rule-learning as individuals or in pairs on subsequent individual rule-learning performance. Individual and paired subjects initially identified four two-value conjunctive rules. Forty-eight hours after the initial learning session, all subjects individually identified four more two-value conjunctive rules from an entirely different array. Both sessions involved the selection paradigm. Bender found no significant differences in the number of trials to solution or total time to solution between individuals and pairs in either session. However, in the learning session, individually trained subjects required less decision time than paired subjects.

Bender's results differed from those of previous researchers, who found groups used fewer trials in the learning session (Laughlin & Sweeney, 1977) and more time in the transfer session (Lemke et al., 1969). These differences may have been due, in part, to differences in the selection paradigm procedure for identifying solved rules. Laughlin & Sweeney used a procedure in which subjects selected an instance from the array, the experimenter informed them whether the instance was correct, subjects made a hypothesis concerning the identity of the rule, and the experimenter informed the subjects whether the rule was correct. Lemke et al. (1969) used the same procedure but deleted the steps involving subjects hypothesizing rules and receiving feedback. Subjects had to state a "conclusion" to try to solve a problem. Bender used the same procedure as Laughlin & Sweeney, but deleted the portion involving feedback to subjects concerning their hypothesized rules. Subjects were also required to state a "conclusion" to solve the problems.

Reasons for the reanalysis are related to the design and analysis of the Bender (1980) study. First, Bender used a multivariate repeated measures design. A multivariate analysis was appropriate, assuming the dependent measures were intercorrelated. However, a repeated measures multivariate analysis may not have been statistically valid. "It seems that analyses for doubly multivariate data have so far been developed only under the assumption that the covariance structure over occasions is identical (in the population) for all variables" (Tatsuoka, 1974, p. 288). Thus, any change in correlation between Problem 1 and Problem 2, and Problem 1 and Problem 3 must be the same for all dependent measures. This assumption of no differential transfer for various dependent measures may be impossible to meet. Therefore, each dependent measure in Experiment 1 is analyzed separately. Second, Bender stressed the use of process and performance measures. Trials to solution is a general performance measure, and decision time is a measure of processing latency. However, two measures of processing efficiency, labeled strategy efficiency and the proportion of untenable hypotheses, were available from Bender's data, but not analyzed. The efficiency of hypothesis-testing is reflected by the strategy efficiency, and the cognitive monitoring of feedback is reflected by the proportion of untenable hypotheses. Therefore, trials to solution, the proportion of untenable hypotheses, strategy efficiency, and decision time are all analyzed in Experiment 1.

This reanalysis provides implications concerning power and the results of Experiment 1, as compared with those of Bender's (1980) study. When the same sample size is used in all analyses, the univariate analyses, by virtue of having fewer dependent variables, have more power than the multivariate analysis. The probability of not rejecting a true result is greater for the univariate analyses. Therefore, results which were not significant in Bender's original analysis may be significant in Experiment 1. Furthermore, previous researchers have only analyzed single dependent measures. The reanalysis should allow a more direct comparison of results between this and previous research.

Research Hypotheses

Results of Experiment 1 must be considered exploratory, since Bender (1980) used a different procedure than previous researchers. Trials to solution and decision time were expected to manifest results similar to those of the original statistical analysis. Previously grouped subjects were expected to decrease, and individually trained subjects to increase, their decision time from the learning session to the transfer session. The direction of change for trials to solution in the original analysis suggested that previously grouped subjects would use more trials to solution in the transfer session than would individually trained subjects.

The proportion of untenable hypotheses and strategy efficiency were expected to evince results similar to those of previous researchers

(Laughlin, 1965). Pairs were expected to use more efficient strategies and a lower proportion of untenable hypotheses than individuals, in the learning session. However, in the transfer session, previously paired subjects were expected to use more untenable hypotheses and less strategy efficiency than individually trained subjects.

Method

Subjects

Subjects were 60 male and female undergraduates from Iowa State University. Each subject was randomly assigned to one of two groups, cooperative pairs or individuals. The individually trained group contained 20 subjects, 10 male and 10 female. The paired group contained 20 pairs of subjects, 10 male and 10 female.

Apparatus

Two stimulus arrays were used in Experiment 1. Each array was constructed on a 33 cm x 69 cm sheet of poster board. The forms array consisted of the dimensions of form, size, and color, with three attributes of each. The cars array consisted of the dimensions of price, resale value, and gas mileage ranges, with three attributes each. Both arrays contained 27 different instances. Time was measured by a Heuer Chronometer, accurate to the tenth of the second.

Procedure

Experiment 1 consisted of two sessions; a learning session, and a transfer session conducted 48 hours later. Four conjunctive rule-learning problems were solved by all subjects in each session. Rules for each session were chosen at random before the formal beginning of the experiment and were randomly arranged before being presented to each pair or individual.

The same general procedure was used for all subjects. The forms array was displayed and described for the subjects. Three categories of rules; labeled single-value, two-value conjunctive, and two-value disjunctive; were defined with examples. Subjects were informed they were to identify four rules, which could be any of the three types. In reality, only two-value conjunctives were used, to make the task more simple. For each problem, one instance which belonged to the set defined by the rule to be discovered was provided for the subjects. Subjects were then asked to state a rule which they considered to be possible, and to select another instance about which they wanted information. Subjects were informed whether or not their instance was in the set defined by the rule to be identified. This sequence of possible rule-instance-feedback was repeated until subjects indicated they thought they knew the identity of the rule. If their conclusion was correct, the problem was solved. If they were incorrect, the conclusion was treated as a possible rule and another instance was tested. The feedback sequence was then continued until the correct rule was discovered. Time was recorded from the completion of stating the initial instance to the statement of the correct rule for each problem. The specific instructions for subjects working in pairs in the learning session are found in Appendix A.

Instruction for the transfer session described the cars array. The types of rules were reviewed, with examples from the cars array. Finally, the procedure was briefly reviewed. All subjects were administered the transfer problems as individuals. Scores of the members of each pair were combined in the transfer session to make use of all the data.

Design

The overall design was a two (individual vs. pair) x two (sex) x two (learning vs. transfer session) factorial, with the last variable treated as a within-subject variable. Four dependent variables, including trials to solution, the proportion of untenable hypotheses, strategy efficiency, and decision time, were measured. A secondary design, a two (individual vs. pair) x two (sex) x four (problems) factorial, with the last variable treated as a within-subjects variable, was used to investigate performance during the learning and transfer sessions. The four dependent variables were measured for each session. The overall research design is presented in Figure 1.

Group	N	Learning Task	Transfer Task
Pairs			
Male	10	Cooperative Pair	Mean Combined Scores
Female	10	Cooperative Pair	Mean Combined Scores
Individuals			
Male	10	Individual	Individual
Female	10	Individual	Individual

Figure 1. Design of Experiment 1

Scoring procedures

Dependent measures included trials to solution, the proportion of untenable hypotheses, strategy efficiency, and decision time. Scores for each session, and problems within sessions were computed for each dependent measure. For trials to solution, each rule-selection-feedback sequence was treated as one trial. The total number of all trials per problem within one session constituted the trials per session score. The trials per transfer session score for previously paired subjects was the mean of their individual trials per session scores.

Untenable hypotheses were scored by the following three rules. First, any rule which described a previous negative instance was counted as one untenable hypothesis. Second, any rule which did not describe a previous positive instance was counted as one untenable hypothesis. Third, the total number of untenable hypotheses per problem was divided by the number of trials for that problem, providing a proportion of untenable hypotheses ranging from 0.00 to 1.00. For individual subjects, and paired subjects during the learning session, the total number of untenable hypotheses for all problems per session was divided by the total number of trials for that session. For previously paired subjects, during the transfer session, the total number of untenable hypotheses for both subjects was divided by the total trials to solution of both.

For strategy efficiency, an instance was counted as efficient if it followed a logically possible rule and eliminated one or more previously possible rules. The total number of efficient selections was divided by the trials per problem, providing a score of strategy efficiency ranging from 0.00 to 1.00. For individual subjects, and paired subjects during the learning session, the total number of efficient selections for all problems per session was divided by the total number of trials for that session. For previously paired subjects, during the transfer session, the total number of efficient selections for both subjects was divided by the total trials to solution of both.

Decision time was measured by dividing the time to solution per problem by the number of trials to solution for each problem. The total time to solution for all problems was divided by the total number of trials for each session to provide a decision time per session. For the decision time of previously paired subjects in the transfer task, the times to solution for both subjects were divided by the total trials for both.

Results

Correlations between problems and between dependent measures, within each session, were computed. Changes in the correlations between problems were measured by a test of differences between dependent correlations (Cohen & Cohen, 1975). These changes were not the same for each dependent measure. Trials to solution was positively correlated with the proportion of untenable hypotheses, and negatively correlated with strategy efficiency, in both sessions. Strategy efficiency was negatively correlated with the proportion of untenable hypotheses, in both sessions.

Learning session

Data from the learning session were analyzed by four separate two (individual vs. pair) x two (sex) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each dependent measure. Tukey's HSDs were applied for multiple comparisons of means. ANOVA summary tables for these analyses are found in Appendix B. Means for each dependent measure on each problem are found in Table 1.

Table 1. Means of all dependent measure for problems in the learning session of Experiment 1[a]

Measures	Problem 1	Problem 2	Problem 3	Problem 4
Trials to solution	4.775	4.850	3.475	3.825
Untenable hypotheses*	0.682	0.842	0.472	0.406
Strategy efficiency**	1.716	1.662	2.039	2.090
Decision time***	45.175	30.225	27.150	24.700

[a]Untenable hypotheses and strategy efficiency are arcsin transformations.

*$p < 0.02$.

**$p < 0.0009$.

***$p < 0.0001$.

Trials to solution No significant main effects or interactions were obtained for the number of trials to solution during the learning session.

Untenable hypotheses The within-subjects main effect of problems was significant $F(3.108) = 3.50$, $p < 0.02$. Problem 2 required a greater proportion of untenable hypotheses than Problem 4, HSD = 0.398, $p < 0.05$.

Strategy efficiency The within-subjects main effect of problems was significant $F(3,108) = 5.99$, $p < 0.0009$. Problem 4 resulted in a greater proportion of efficient selections than Problem 1, HSD = 0.335, $p < 0.05$, and Problem 2, HSD = 0.411, $p < 0.01$. Problem 3 resulted in a greater proportion of efficient selections than Problem 2, HSD = 0.335, $p < 0.05$.

Decision time The between-subjects main effect of group size was significant $F(1,36) = 5.28$, $p < 0.03$, with mean decision times for individuals and pairs of 26.95 and 36.675 seconds, respectively. The within-subjects main effect of problems was also significant $F(3,108) = 18.83$, $p < 0.0001$. Problem 1 required more decision time than any other problem, HSD = 9.72, $p < 0.01$.

Transfer session

Data from the transfer session were analyzed by four separate two (individual vs. pair) x two (sex) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each dependent measure. Tukey's HSDs were performed as tests of multiple comparisons of means. Summary tables for these ANOVAs are found in Appendix B. Means for each dependent measure on each problem are found in Table 2.

Trials to solution No main effects or interactions were significant for trials to solution during the tranfer session.

Untenable hypotheses The between-subjects main effect of group size was significant $F(1,36) = 16.38$, $p < 0.0003$, with mean arcsin transformed proportions of untenable hypotheses for subjects initially trained as pairs and individuals of 1.2232 and 0.6343, respectively. The within-subjects main effect of problems was also significant $F(3,108) = 2.75$, $p < 0.05$. However, no comparison between problem means was significant.

Table 2. Means of all dependent measures for problems in the transfer session of Experiment 1[a].

Measures	Problem			
	1	2	3	4
Trials to solution	4.675	4.350	3.900	3.325
Untenable hypotheses*	1.141	1.082	0.787	0.705
Strategy efficiency**	1.484	1.509	1.706	1.971
Decision time***	43.175	30.925	25.700	27.500

[a]Untenable hypotheses and strategy efficiency are arcsin transformations.

*$p < 0.05$.

**$p < 0.001$

***$p < 0.0001$.

Strategy efficiency The between-subjects main effect for group size was significant $F(1,36) = 8.95$, $p < 0.005$, with the mean transformed proportions of efficient selections for subjects initially trained as pairs and individuals of 1.4419 and 1.8933, respectively. The within-subjects main effect of problems was also significant $F(3,108) = 5.93$,

$p < 0.001$. Subjects were more efficient on Problem 4 than on Problems 1 and 2, HSD = 0.4251, $p < 0.01$.

Decision time The within-subjects main effect on problems was significant $F(3,108) = 23.61$, $p < 0.0001$. Problem 1 required more decision time than Problems 2 through 4, HSD = 7.4345, $p < 0.01$.

The two-way interaction of group size x problems was also significant $F(3,108) = 4.62$, $p < 0.005$. Means for this interaction are found in Table 3. Subjects initially trained in pairs required more decision time for Problem 1 than Problems 2 through 4, and than any problems solved by subjects initially trained as individuals, HSD = 7.4345, $p < 0.01$.

Table 3. Means of decision time for group size x problems in the transfer session of Experiment 1

Group Size	Problem			
	1	2	3	4
Individuals	38.00[a]	33.60[b]	24.95	29.10
Pairs	48.35[c]	28.25	26.45	25.90

[a]Significantly more than Individual's 3 or 4 and Pair's 2 through 4 at $p < 0.01$.

[b]Significantly more than Individual's and Pair's 4 at $p < 0.01$.

[c]Significantly more than any other problem at $p < 0.01$.

Individually trained subjects required more decision time for Problem 1 than for Problems 3 or 4, and than subjects trained as pairs required for Problems 2 through 4, HSD = 7.4345, $p < 0.01$. Finally, individually trained subjects required more decision time for Problem 2 than Problem 3, and than subjects trained as pairs required for Problem 4, HSD = 7.4345, $p < 0.01$.

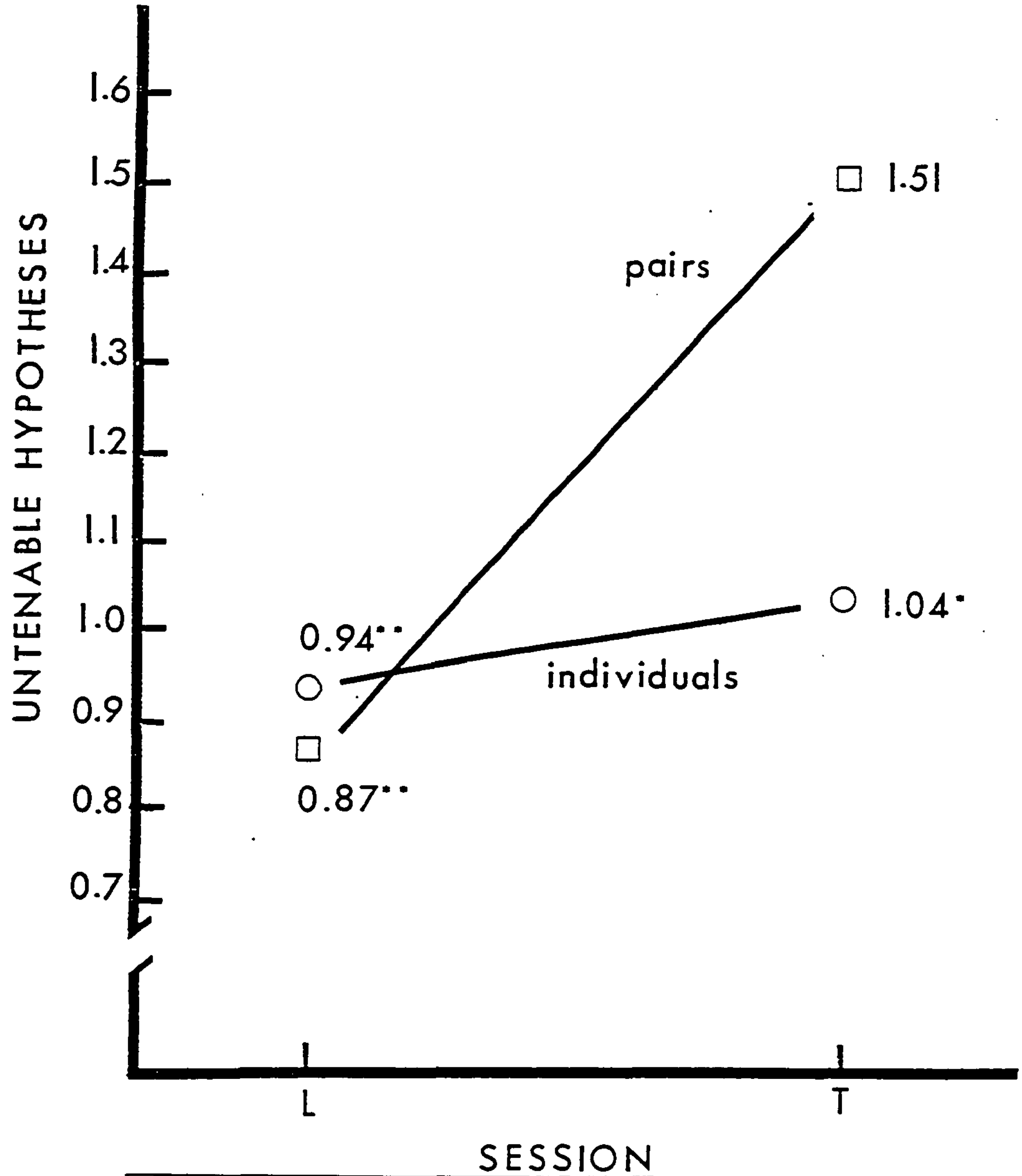
Transfer across sessions

The analyses of data reflecting changes in performance from the learning to the transfer session consisted of four separate two (individual vs. pair) x two (sex) x two (learning vs. transfer session) split-plot ANOVAs, with sessions treated as a within-subjects variable. One ANOVA was completed for each dependent measure. Tukey's HSDs were performed as multiple comparisons of means. Summary tables for these ANOVAs are found in Appendix B.

Trials to solution No significant main effects or interaction were obtained for trials to solution.

Untenable hypotheses The within-subjects main effect of sessions was significant $F(1,36) = 11.97$, $p < 0.002$, with mean transformed proportions of untenable hypotheses for the learning and transfer sessions of 0.9048 and 1.2724, respectively.

The two-way interaction of group size x sessions was also significant $F(1,36) = 6.23$, $p < 0.02$. Means for this interaction are found in Figure 2. Subjects initially trained in pairs used a higher proportion of untenable hypotheses in the transfer session than subjects from either group size required during the learning session, HSD = 0.5100, $p < 0.01$. Subjects initially trained in pairs also used a greater proportion of untenable hypotheses in the transfer session than did subjects initially trained as individuals, HSD = 0.4091, $p < 0.05$.



• less than pairs at T at 0.05
•• less than pairs at T at 0.01

FIGURE 2 Mean Untenable Hypotheses for Group Size x Session in Experiment I.

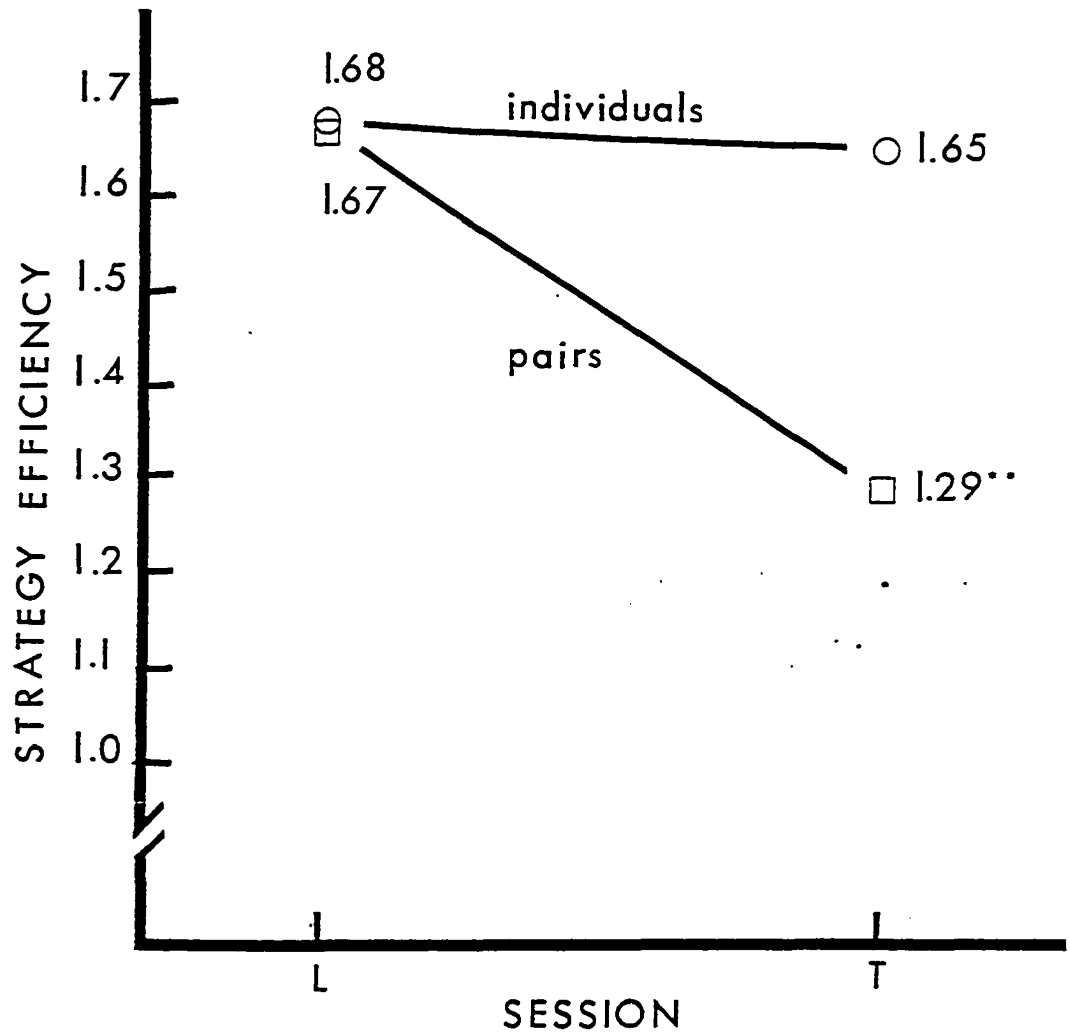
<u>Strategy efficiency</u> The within-subjects main effect of sessions was significant $F(1,36) = 9.29$, $p < 0.005$, with mean transformed proportions for the learning and transfer sessions of 1.6741 and 1.4676, respectively.

The two-way interaction of group size x sessions was also significant $F(1,36) = 6.80$, $p < 0.02$. Means for this interaction are depicted in Figure 3. Subjects initially trained as pairs used a lower proportion of efficient selections in the transfer session than individually trained subjects used in either session, and than paired subjects used during the learning session, HSD = 0.3253, $p < 0.01$.

<u>Decision time</u> The two-way interaction of group size x sessions was significant $F(1,36) = 6.00$, $p < 0.02$. Means for this interaction are found in Figure 4. In the learning session, pairs used more decision time than did individuals, HSD = 8.717, $p < 0.01$.

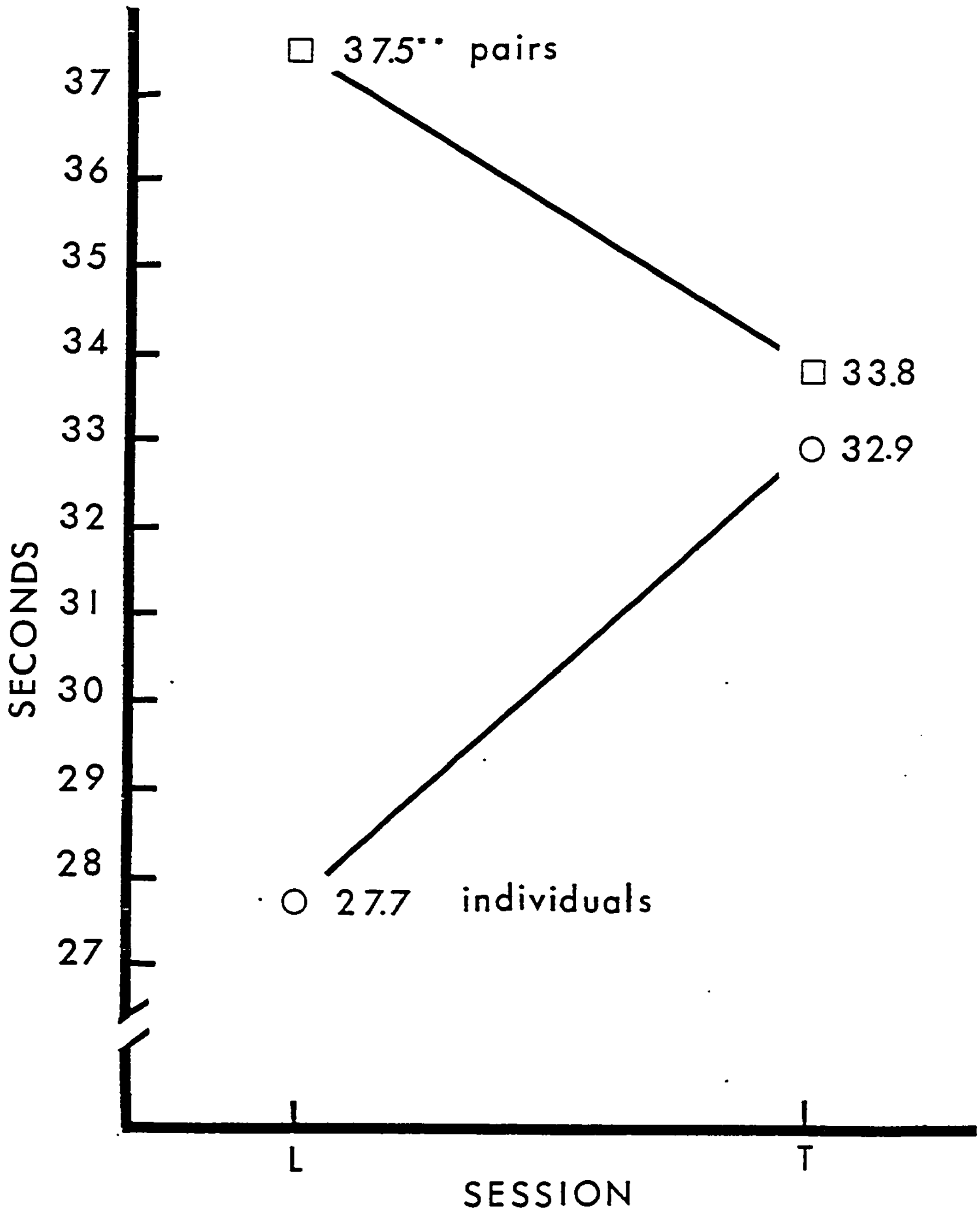
Discussion

Results of Experiment 1 were expected to be similar to the results of previous research. In the learning session, pairs were expected to require more decision time, a lower proportion of untenable hypotheses, and a greater strategy efficiency than individuals. Transfer to the second session was expected to result in more trials to solution, a greater proportion of untenable hypotheses, and less strategy efficiency for previously paired subjects than for individually trained subjects. These expectations were only partially fulfilled.



** less than any other at 0.01

FIGURE 3. Mean Strategy Efficiencies for Group Size x Session in Experiment I



** more than individuals at L at 0.01

FIGURE 4. Mean Decision Times for Group Size x Sessions in Experiment I

Metacognitively, working with a partner, in this experiment, was an external variable which tended to negatively affect the metacognitive control of subsequent individual rule-learning. Paired subjects required more decision time than individuals, in the learning session. In the transfer session, previously paired subjects required a greater proportion of untenable hypotheses and used a lower strategy efficiency when tested individually than when working as a pair. These same subjects also used a greater proportion of untenable hypotheses and a lower strategy efficiency than individually trained subjects in the transfer session. Apparently, individually trained subjects developed metacognitive knowledge which positively transferred from the learning to the transfer sessions; whereas, paired subjects did not individually process the information in the transfer session as well as they did as pairs in the learning session.

Transfer on successive problems

Although trials to solution, the proportion of untenable hypotheses, and strategy efficiency were intercorrelated, differential transfer did occur. Thus, the use of separate repeated measures ANOVAs was the appropriate analysis.

Transfer from problem to problem within either session was near transfer. No positive near transfer was evident for trials to solution in either session. However, the proportion of untenable hypotheses, strategy efficiency, and decision time exhibited positive near transfer to some degree in one or both sessions.

Learning session The time required for decision-making stabilized early, while improvements in the interpretation of feedback and testing of instances continued to occur. Problem 1 required the most decision time of any problem, and less strategy efficiency than Problems 3 or 4. Problem 2 resulted in a greater proportion of untenable hypotheses than Problem 4, and less strategy efficiency than Problem 3. Positive near transfer occurred for measures of processing latency and efficiency, but not for the measure of general performance.

Transfer session As in the learning session, the pattern of stabilization of decision time, followed by improvement in processing, was evident. Problem 1 required more decision time than any other problem. Positive transfer was not evident in the proportion of untenable hypotheses. However, subjects used less strategy efficiency for Problems 1 and 2 than Problem 4. Positive near transfer was reflected by procesing time and the efficiency of hypothesis-testing, but not by the interpretation of feedback or the general performance.

Transfer across sessions

Contrary to the expected results, pairs did not use better processing in the learning session; however, the predicted decline in processing for previously paired subjects in the transfer session did occur. In the learning session, the performance of pairs and individuals were similar, with the exception that pairs required more decision time. Presumably, discussion by the members of the pair was responsible for the increased time. In the transfer session, no differences occurred between groups in processing time or general performance; however, previously paired subjects evinced a decline in the processing of feedback and efficiency of hypothesis-testing. These results indicate that rule-learning with a partner does not always improve performance, and may lead to poorer processing when subsequently rule-learning as individuals.

Conclusion

Two major conclusions are readily apparent from the results of Experiment 1. First, although previous research has indicated small groups tend to outperform individuals in rule-learning tasks, this is not always true. Pairs and individuals used approximately the same amounts of trials to solution, untenable hypotheses, and strategy efficiency in the learning session. Second, rule-learning with a partner did not improve subsequent individual rule-learning performance. In fact, subjects who had been previously paired evinced a decline in processing performance when tested individually. Previously paired subjects used a higher proportion of untenable hypotheses and a lower strategy efficiency in the transfer session, than in the learning session. Individually trained subjects did not evince this decline, and used a lower proportion of untenable hypotheses and a higher strategy efficiency in the transfer session than did previously paired subjects.

Unfortunately, this experiment confounds any differences in the difficulty of the forms and cars arrays with the sessions. The forms array may have been so easy for both pairs and individuals that both groups performed near the same level. If a transfer array which was less difficult than the cars array had been used, the performance of previously paired and individually trained subjects may have been different.

One method of separating the effects of sessions from the difficulty of the arrays would be to counterbalance the presentation of the arrays. Thus, one group of subjects would be tested using the forms then cars arrays; and another, the cars then forms arrays. This order of presentation could also be included in the analysis. The use of counterbalancing to separate the effects of arrays from those of sessions is introduced in Experiment 2. Two new arrays, designed to be similar to the forms and cars arrays, are also included and counterbalanced. Experiment 2 is designed primarily as a test of the various arrays and the counterbalancing procedure. Therefore, group size is not included; however, sex remained a fixed variable.

EXPERIMENT TWO

Experiment 2 was designed to promote the generalizability of results concerning rule-learning performance of indiviudal subjects. In order to promote generalizability, two additional arrays, called borders and houses, were developed for use with the forms and cars arrays. The forms and borders arrays were perceptually oriented, while the cars and houses arrays were more semantically oriented. In order to investigate transfer from the learning session to the delayed transfer session, without the effects of group size, only individual subjects were used.

Independent variables of interest included the general presentation order of the arrays and sex of the subject. Counterbalancing the presentation order led to two general presentation orders, perceptual-to-semantic and semantic-to-perceptual. Since no significant results involving sex were found in Experiment 1, sex was also included, to discover if any significant results occur with a variety of stimuli.

Research Hypotheses

Experiment 2 was exploratory with regard to the performances elicited by the various arrays and general presentation orders. Therefore, only the null hypotheses of no differences between the arrays or presentation orders were offered. No difference in performance between the sexes was expected. Finally, the performance of individuals in the transfer session was expected to be superior to performance in the learning session.

Method

Subjects

Subjects were 32 male and female undergraduates from Iowa State University. Each subject was randomly assigned to one of eight conditions, with the constraint of equal representation of the sexes within each condition. Each condition contained four subjects, two male and two females. One male had to be replaced, as his trials to solution score for the learning session was over three standard deviations above the mean for all subjects.

Apparatus

The forms array consisted of three dimensions, with three values each. Dimensions were geometric shape, color and size. Shapes included squares, triangles and circles; colors included red, yellow and blue; and sizes included large, medium and small. The 27 different combinations of these values composed the instances of the array. An example of a conjunctive rule from this array was, "large and yellow objects."

The cars array was more semantically-oriented than the forms array. The three dimensions of price, resale value and gas mileage, had three values of high, medium and low. Each dimension and value combination was described by initials and numbers. For example, "medium gas mileage" was represented by "MG 19-27." The cars array also contained 27 instances. An example of a conjunctive rule from this array was "high price and medium gas."

The borders array was designed to be perceptually oriented, similarly to the forms array. The array consisted of 27 rectangles containing one value from each of three dimensions. The three dimensions included the number of borders around the rectangle, the shape within the rectangle, and the number of shapes. The number of borders ranged from zero to two. Shapes were hexagons, arrows and stars. The number of shapes ranged from one to three. An example of a conjunctive rule from this array was, "no borders and two shapes."

The houses array was another semantically-oriented array. It described 27 houses, on the basis of style, location and number of bedrooms. Styles were ranch, A-frame and condo. Locations included city, suburb and country. The number of bedrooms ranged from two to four. An example of a conjunctive rule from this array was, "four bedrooms and country."

Each stimulus array was constructed on a 33 cm x 69 cm sheet of poster board. Time was measured by a Northeast Instruments stopwatch, accurate to the tenth of the second.

Procedure

The experiment consisted of two sessions: a learning session, and a transfer session conducted 48 hours later. Four two-value conjunctive rule-learning problems were solved by the subjects in each session. Rules for each session were determined at random before the formal beginning of the experiment, and were randomly arranged before being presented to each subject.

Instructions for the learning session first provided a general description of the task. This was followed by the presentation and description of the array for that condition. Single-value, two-value conjunctive, and two-value disjunctive rules were defined. An example of each was provided. Subjects were informed that they were to identify four rules, which may have been any of the three types. In reality, only two-value conjunctive rules were used.

A simple procedure comprised the identification task for each rule. First, an instance described by the rule to be identified was provided for the subjects. Subjects were informed that the instance belonged in the set defined by the rule. Subjects then stated a possible rule; that is, a rule which the subjects thought could have been the rule to be identified. No feedback was provided for these hypothesized rules. Subjects then chose an instance from the array and were informed whether that instance belong in the set defined by the rule to be identified. The sequence of stating a rule, selecting an instance and receiving feedback was continued until subjects thought they knew the identity of the rule. Subjects then stated the rule, as a conclusion. If the conclusion was correct, the problem was solved. If it was incorrect, the identification sequence was repeated until the correct rule was discovered. Subjects were urged to make conclusions only after they felt they were positive of the identity of the rule.

Instructions for the transfer session described the new array. The types of rules were reviewed, with examples from the new array. Finally, the procedure was briefly reviewed.

Counterbalancing the presentation order across sessions resulted in two general presentation orders. The semantic-to-perceptual order included the conditions of cars-to-forms, cars-to-borders, houses-to-forms, and houses-to-borders arrays. The perceptual-to-semantic order included the conditions of forms-to-cars, forms-to-houses, borders-to-cars, and borders-to-houses arrays.

Design

The overall design was a two (sex) x two (grouped order: perceptual-to-semantic vs. semantic-to-perceptual) x two (session) factorial, with the last variable treated as a within-subjects variable. Four dependent measures of trials to solution, the proportion of untenable hypotheses, strategy efficiency, and decision time were used. A second design, a two (sex) x two (grouped order) x four (problems) factorial, with problems treated as a within-subjects variable, was used to investigate performance during each session. The four dependent measures were used in each session. The overall design is found in Figure 5.

The four dependent measures were measured by the same procedures used for individual subjects in Experiment 1. The success of counterbalancing was determined by performing *t*-tests between means for each of two conditions using the same initial array. The initial difficulty of each array was determined by comparing the dependent measures for each array in the learning session. The difficulty of each array following training on a different style of array was investigated by comparing the dependent measures for each array in the transfer session.

Group	N	Learning Session	Transfer Session
Perceptual-to-semantic			
Males	8	Individuals	Individuals
Females	8	Individuals	Individuals
Semantic-to Perceptual			
Males	8	Individuals	Individuals
Females	8	Individuals	Individuals

Figure 5. Design of Experiment 2

Results

The success of counterbalancing the various arrays was tested by *t*-tests of the difference between means. These *t*-tests indicated no significant differences. Correlations between problems and between dependent measures, within each session, were also computed. Tests of differences between dependent correlations (Cohen & Cohen, 1975) indicated the existence of differential transfer from problem to problem for the dependent measures. The proportion of untenable hypotheses was positively correlated with the number of trials to solution, and negatively correlated with strategy efficiency, in both sessions. Trials to solution and strategy efficiency were negatively correlated in the learning session.

Learning session

Data from the learning session were analyzed by four separate two (sex) x four (arrays) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each of the four dependent measures. Tukey's HSDs were used as tests of multiple comparisons of means. ANOVA summary tables for each analysis are found in Appendix B. Means for each dependent measure on each problem are found in Table 4.

Table 4. Means of all dependent measures for problems in the learning session of Experiment 2[a]

Measures	Problem			
	1	2	3	4
Trials to solution	8.563[b]	6.094	5.250	4.563
Untenable hypotheses	1.269	1.264	0.994	0.926
Strategy efficiency	1.388	1.534	1.516	1.768
Decision time	45.250[c]	26.969	23.844	21.969

[a]Untenable hypotheses and strategy efficiency are arcsin transformations.

[b]Significantly more than Problem 3 at $p < 0.05$, and Problem 4 at $p < 0.004$.

[c]Significantly more than any other Problem at $p < 0.0001$.

Trials to solution The within-subjects main effect of problems was significant $F(3,72) = 5.05$, $p < 0.004$. Subjects required more trials to solution for Problem 1 than for Problem 3, HSD = 2.9076, $p < 0.05$; and Problem 4, HSD = 3.5683, $p < 0.01$.

Untenable hypotheses No significant main effects or interactions were obtained for the proportion of untenable hypotheses, during the learning session.

Strategy efficiency No significant main effects or interaction were obtained for strategy efficiency, during the learning session.

Decision time The between-subjects main effect of arrays was significant $F(3,24) = 3.11$, $p < 0.05$. Table 5 displays the mean decision times of all arrays used in the learning session. No significant comparisons between means were obtained. The within-subjects main effect of problems was significant $F(3,72) = 19.78$, $p < 0.0001$. Problem 1 required more decision time than any other problem, HSD = 11.4566, $p < 0.01$.

Table 5. Mean decision times for each array in the learning session of Experiment 2

Arrays	Time
Forms	24.250
Cars	34.000
Borders	33.938
Houses	24.219

The two-way interaction of arrays x problems was significant $F(9,72) = 2.50$, $p < 0.02$. Table 6 displays the means of this interaction. Problem 1 of the borders array required more decision time than any other problem, except Problem 1 of the cars array, HSD = 29.1533, $p < 0.01$. Problem 1 of the cars array required more decision time than Problem 4 of the forms, borders and house arrays HSD = 29.1533, $p < 0.01$. Problem 1 of the cars array also required more decision time than Problems 2 or 3 of the forms array and Problem 3 of the houses array, HSD = 25.2544, $p < 0.05$.

Table 6. Mean decision times for arrays x problems in the learning session of Experiment 2

Arrays	Problem 1	Problem 2	Problem 3	Problem 4
Forms	34.875	21.375	22.750	18.000
Cars	48.250[a]	34.625	25.125	28.000
Borders	66.375[b]	25.250	26.375	17.750
Houses	31.500	26.625	21.125	17.625

[a]Significantly more than Forms 4, Borders 4 and Houses 4 at $p < 0.01$; and more than Forms 2 or 3, and Houses 3 at $p < 0.05$.

[b]Significantly more than any other Problem, except Cars 1, at $p < 0.01$.

Transfer session

Data from the transfer session were analyzed by four separate two (sex) x four (arrays) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each dependent measure. Tukey's HSDs were used as tests of multiple comparisons of means. ANOVA summary tables for each analysis are found in Appendix B.

Trials to solution No significant main effects or interactions were obtained for trials to solution, during the transfer session.

Untentable hypotheses No significant main effects or interactions were obtained for the proportion of untenable hypotheses, during the transfer session.

Strategy efficiency No significant main effects or interactions were obtained for strategy efficiency, during the transfer session.

Decision time The between-subjects main effect of array was significant $F(3,24) = 3.42$, $p < 0.04$. The cars array required more decision time than the houses array, HSD = 12.1756, $p < 0.05$. Table 7 presents the mean decision times for each array.

Table 7. Mean decision times for each array in the transfer session of Experiment 2

Arrays	Time
Forms	19.906
Cars	31.781[a]
Borders	23.188
Houses	19.219

[a]Significantly more than the houses array at $p < 0.05$.

Transfer across sessions

The analyses of data reflecting changes in performance from the learning to transfer sessions consisted of four separate two (sex) x two (grouped order: semantic-to-perceptual vs. perceptual-to-semantic) x two (learning vs. transfer session) split-plot ANOVAs, with sessions treated as a within-subjects variable. One ANOVA was completed for each dependent measure. Tukey's HSDs were used to test multiple comparison of means. Summary tables of all ANOVAs are found in Appendix B.

Trials to solution The within-subjects main effect of sessions was significant $F(1,28) = 18.15$, $p < 0.0002$, with mean trials for the learning and transfer sessions of 24.4688 and 17.5313, respectively.

Untenable hypotheses The within-subjects main effect of sessions was significant $F(1,28) = 6.03$, $p < 0.03$, with mean transformed proportions of untenable hypotheses for the learning and transfer sessions of 1.4277 and 1.1573, respectively.

Strategy efficiency No main effects or interactions were significant for strategy efficiency.

Decision time The within-subjects main effects of sessions was significant $F(1,28) = 13.99$, $p < 0.0008$, with mean decision times for the learning and transfer sessions of 32.5938 and 24.5313 seconds, respectively.

Discussion

The experience of solving a series of rule-learning problems led to improved performance on similar problems. This improvement was evident for transfer from problem to problem, and from a learning to a delayed transfer session. Results were not affected by the sex of the subjects, and occurred over a variety of stimulus arrays. Furthermore, it made no significant difference whether the initial task was perceptually oriented and followed by a more semantic task, or semantically oriented and followed by a more perceptual task.

Solving rule-learning problems was an experience which, theoretically, led to more positive megacognitive control of subsequent rule-learning. Subjects used fewer trials to solution, less time for evaluating and obtaining information, and better interpretation of information in the transfer session than in the learning session. The problem-solving process evincing the most improvement was the interpretation and use of feedback, rather than the obtaining of the information through efficient hypothesis-testing.

Transfer on successive problems

Near transfer was reflected by trials to solution and decision time, in the learning session. No near transfer was reflected by any dependent measure in the transfer session.

Learning session   Positive near transfer was exhibited by trials to solution and decision time. Problem 1 required more trials than Problems 3 or 4. A significant decision time main effect of problems was probably due to Problem 1 of the borders and cars arrays. Problem 1 of the borders array required more decision time than any other problem, except Problem 1 of the cars array. Problem 1 of the cars array required more decision time than Problems 2 through 4 of the cars array, than Problems 3 or 4 of the houses array, and than Problem 4 of the borders array. Individuals appeared to improve in both the trials to solution and decision time from Problem 1. However, the process measures of the strategies used by these individuals did not reflect improvement. Apparently, subjects employed their strategies more effectively, and with less decision time, without improving the processing itself.

Transfer session No near transfer was displayed by any of the four dependent measures; however, the decision time main effect for arrays was significant. The cars array required more decision time than the houses array.

Transfer across sessions

Positive transfer from the learning to the transfer session was reflected by the number of trials to solution, the proportion of untenable hypotheses and decision time. Subjects used fewer trials in solving the transfer problems, made better use of feedback in determining possible rules, and required less time for obtaining and processing information, in the transfer task. This transfer occurred across both the presentation order and sex.

Conclusion

Individuals exhibited positive transfer of rule-learning performance both within the learning session and from the learning to the delayed transfer session. The within sessions transfer did not reflect improvement in processing, whereas the between sessions transfer reflected improvement in the interpretation of feedback. This positive transfer occurred across sex and a variety of arrays.

EXPERIMENT THREE

Experiment 3 was designed to investigate the effects of vocalization and small group rule-learning on the individual transfer of performance. Variables included vocalization of reasoning, group size and the presentation order of the grouped arrays. Since the cars array required more decision time than did any other array in Experiment 2, all four arrays and eight combinations were used. Thus, the difficulty of the various arrays, and perceptual-to-semantic vs. semantic-to-perceptual transfer, were also investigated.

This experiment differed from Durling & Schick (1976) in four aspects. Implicit in the Durling & Schick study was the assumption that discussion always involved the vocalization of reasoning. However, simply allowing group discussion may not have guaranteed that the reasoning was stated or discussed. Therefore, Durling & Schick separated discussion from nondiscussion, but did not separate vocalization from discussion. To do so, Experiment 3 was a comparison of the rule-learning performance of discussing pairs of subjects who were asked to state their reasoning, discussing pairs of subjects who were not asked to state their reasoning, individuals who were asked to state their reasoning, and individuals who were not required to state their reasoning. A second difference was that Experiment 3 used the same feedback technique as Experiments 1 and 2, not the reduced information feedback technique.

Third, the design of the Durling & Schick study allowed for an investigation of transfer from problem to problem, while Experiment 3 allowed the study of transfer from problem to problem, as well as between sessions. Finally, Durling & Schick used only female subjects, while Experiment 3 involved both males and females.

Research Hypotheses

Vocalizing pairs and individuals were expected to use fewer trials to solution, more efficient strategies, a lower proportion of untenable hypotheses, and more decision time than nonvocalizing pairs and individuals, during the learning session. Second, nonvocalizing pairs were expected to use more decision time than nonvocalizing individuals, in the learning session. Third, in the transfer session, subjects from nonvocalizing pairs were expected to use more trials to solution than nonvocalizing individuals, who were expected to use more than subjects from vocalizing pairs or vocalizing individuals. Fourth, in the transfer session, subjects from vocalizing conditions, who were expected to use more efficient strategies than nonvocalizing individuals, who were expected to be more efficient than subjects from nonvocalizing pairs. Fifth, in the transfer session, subjects from nonvocalizing pairs were expected to use a greater proportion of untenable hypotheses than nonvocalizing individuals, who were expected to use more than any vocalizing subject. Sixth, subjects from nonvocalizing pairs were expected to use more decision time than nonvocalizing individuals, who were expected to use more than any vocalizing subject, in the transfer session. Finally, no differential transfer from the learning session to the transfer session was expected for the types of arrays.

Method

Subjects

Subjects were 96 male and female undergraduates from Iowa State University. Subjects were randomly assigned to one of four groups. The groups were composed of cooperative pairs or individuals, who were either required or not required to vocalize their reasoning during the first session of rule-learning. Groups contained an equal number of males and females. Individually trained groups contained 16 subjects each. Paired groups contained 16 pairs of subjects each, eight male pairs and eight female pairs. One pair of females had to be replaced, as their trials to solution score, in the transfer session, was over three standard deviations above the mean of the rest of the pairs in that vocalization condition.

Apparatus

The arrays were identical to those used in Experiment 2. Time was measured by a Northeast Instruments stopwatch, accurate to the tenth of the second.

Procedure

The experiment consisted of two tasks: one learning, and one transfer task conducted 48 hours later. Four conjunctive rule-learning problems were solved by the subjects in each task. Rules for each task were chosen at random and randomly arranged before being presented to each individual or pair.

Four groups of subjects were administered the tasks. Two groups performed the initial task as individuals; two, as pairs. All groups performed as individuals in the transfer task. One individually trained and one paired group were asked to vocalize part of their reasoning, during the learning task. Reasons such as, "I don't know" were not allowed. No group was required to vocalize their reasoning during the transfer task; although, spontaneous vocalizations were allowed.

The basic procedure and instructions were similar to those used in Experiments 1 and 2, with some changes for the vocalization groups. The learning session procedure for these groups was changed to the following steps. First, a positive instance was described and a possible rule requested. Second, the subjects selected an instance and received feedback as to whether the instance was in the set defined by the rule to be identified. Third, subjects answered the question, "What does that information tell you about possible rules?" The sequence of selection-feedback-reasoning-hypothesis was repeated until subjects reached a conclusion. If the conclusion was correct, the problem was solved. If the conclusion was incorrect, the identification sequence was repeated until a correct conclusion was reached.

Design

The overall design was a two (individual vs. pair) x two (vocalization vs. nonvocalization) x two (grouped order: perceptual-to-semantic vs. semantic-to-perceptual) x two (learning vs. transfer session) factorial, with the last variable treated as a within-subjects variable. The four dependent measures of trials to solution, the proportion of untenable hypotheses, strategy efficiency, and decision time were scored for each session. Scoring procedures were identical to those used in Experiment 1. A second design, a two (individual vs. pair) x two (vocalization vs. nonvocalization) x two (grouped order) x four (problems) factorial, was used to investigate performance during the learning and transfer sessions. The overall design is presented in Figure 6.

Group	N	Learning Task	Transfer Task
Pairs			
Vocalizing	8	Perceptual array	Semantic array
	8	Semantic array	Perceptual array
Nonvocalizing	8	Perceptual array	Semantic array
	8	Semantic array	Perceptual array
Individuals			
Vocalizing	8	Perceptual array	Semantic array
	8	Semantic array	Perceptual array
Nonvocalizing	8	Perceptual array	Semantic array
	8	Semantic array	Perceptual array

Figure 6. Design of Experiment 3

Results

Success of counterbalancing the various arrays was tested by _t_-tests of differences between means. No significant results were obtained. Correlations between problems and between dependent measures, within each session, were also computed. Tests of differences between dependent correlations (Cohen & Cohen, 1975) indicated the existence of differential transfer from problem to problem for the dependent measures. Trials to solution was positively correlated with the proportion of untenable hypotheses, and negatively correlated with strategy efficiency, in both sessions. The proportion of untenable hypotheses and strategy efficiency were negatively correlated, in both sessions.

Learning session

Data from the learning session were analyzed by four separate two (individual vs. pair) x two (vocalization vs. nonvocalization) x four (arrays) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each dependent measure. Tukey's HSDs were used as tests of multiple comparisons among means. ANOVA summary tables for each analysis are found in Appendix B. Means for each dependent measure on each problem are found in Table 8, while the means for each dependent measure on each array are found in Table 9.

Trials to solution The between-subjects main effect of group size was significant $F(1,48) = 6.72$, $p < 0.02$, with mean trials to solution for individuals and pairs of 5.2578 and 4.4063, respectively.

Table 8. Means of all dependent measures for problems in the learning session of Experiment 3

Measures	Problem 1	2	3	4
Trials to solution	6.016[a]	4.672	4.234	4.406
Untenable hypotheses	0.594	0.504	0.482	0.344
Strategy efficiency	1.621[b]	1.791	1.929	1.974
Decision time	56.172[c]	36.016[d]	32.179	28.984

[a] Significantly more than Problems 2 through 4 at $p < 0.01$.

[b] Significantly less than Problems 3 or 4 at $p < 0.01$.

[c] Significantly more than any other Problem at $p < 0.01$.

[d] Significantly more than Problem 4 at $p < 0.01$.

Table 9. Means of all dependent measures for arrays in the learning session of Experiment 3

Measures	Arrays: Forms	Cars	Borders	Houses
Trials to solution	4.281	5.781[a]	4.500	4.766
Untenable hypotheses	0.287	0.853[b]	0.285	0.500
Strategy efficiency	1.890	1.635	2.059[c]	1.730
Decision time	38.625	43.890	34.078	37.297

[a] Significantly more than Forms or Borders at $p < 0.05$.

[b] Significantly more than Forms or Borders at $p < 0.01$.

[c] Significantly more than Cars or Houses at $p < 0.05$.

The between-subjects main effect of arrays was also significant $F(3,48) = 4.08$, $p < 0.02$. The cars array required more trials to solution than the forms or borders arrays, HSD = 1.2449, $p < 0.05$.

The within-subjects main effect of problems was also significant $F(3,144) = 12.55$, $p < 0.0001$. Problem 1 required more trials to solution than Problems 2 through 4, HSD = 1.005, $p < 0.01$.

The two-way interaction of arrays x problem was also significant $F(9,144) = 3.19$, $p < 0.002$. Means for this interaction are found in Table 10. Problem 1 of the cars array required more trials to solution than any other problem in any array, HSD = 2.508, $p < 0.01$.

Table 10. Mean trials to solution for arrays x problems in the learning session of Experiment 3

Arrays	Problems			
	1	2	3	4
Forms	4.625	3.813	3.875	4.813
Cars	8.563[a]	5.375	4.813	4.375
Borders	5.188	4.688	3.938	4.188
Houses	5.688	4.813	4.313	4.250

[a]Significantly more than any other problem of any array at $p < 0.01$.

Untenable hypotheses The between-subjects main effect of group size was significant $F(1,48) = 5.39$, $p < 0.03$, with mean transformed proportions of untenable hypotheses for individuals and pairs of 0.6028 and 0.3598, respectively. The between-subjects main effect of vocalization was significant $F(1,48) = 6.45$, $p < 0.02$, with means for vocalization and nonvocalization of 0.3484 and 0.6243, respectively. Finally, the between-subjects main effect of arrays was significant $F(3,48) = 6.53$, $p < 0.0009$. Subjects used a greater proportion of untenable hypotheses for the cars array than for the forms or borders arrays, HSD = 0.4919, $p < 0.01$.

The two-way interaction of group size x vocalization was also significant $F(3,48) = 5.87$, $p < 0.02$. Means for this interaction are found in Table 11. Individually trained subjects who were not asked to vocalize used a greater proportion of untenable hypotheses than did any other group, HSD = 0.3966, $p < 0.05$.

Table 11. Mean untenable hypotheses for group size x vocalization in the learning session of Experiment 3

Group Size	No	Yes
Individuals	0.863[a]	0.343
Pairs	0.366	0.354

[a]Significantly more than any other group at $p < 0.05$.

Strategy efficiency The between-subjects main effect of group size was significant $F(1,48) = 15.90$, $p < 0.0002$, with mean transformed proportions of efficient selections for individuals and paris of 1.6783 and 1.9791, respectively. The between-subjects main effect arrays was also significant $F(3,48) = 6.10$, $p < 0.002$. Subjects used a greater proportion of efficient selections for the borders array than for the cars or houses arrays, HSD = 0.2959, $p < 0.05$.

The within-subjects main effect of problems was also significant $F(3,144) = 7.12$, $p < 0.0002$. Subjects used a lower proportion of efficient selections for Problem 1 than for Problems 3 or 4, HSD = 0.2613, $p < 0.01$.

The two-way interaction of group size x vocalization was also significant $F(1,48) = 5.22$, $p < 0.03$. Means for this interaction are found in Table 12. Individually trained subjects, who were not required to vocalize, used a lower proportion of efficient selections than non-vocalizing subjects trained in pairs, HSD = 0.3545, $p < 0.01$; and than any vocalizing group, HSD = 0.2821, $p < 0.05$.

Table 12. Mean strategy efficiencies for group size x vocalization in learning session of Experiment 3

Group Size	Vocalization	
	No	Yes
Individuals	1.535[a]	1.822
Pairs	2.008	1.951

[a]Significantly less than nonvocalizing pairs at $p < 0.01$ and either vocalizing group at $p < 0.05$.

<u>Decision time</u> The between-subjects main effect of group size was significant $F(1,48) = 14.60$, $p < 0.0004$, with mean decision times for individuals and pairs of 33.2422 and 43.7031 seconds, respectively. The between-subjects main effect of vocalization was also significant $F(1,48) = 16.73$, $p < 0.0002$, with means for vocalization and nonvocalization of 44.0703 and 32.8750 seconds, respectively.

The within-subjects main effect of problems was also significant $F(3,144) = 72.28$, $p < 0.0001$. Problem 1 required more decision time than any other, and Problem 2 required more than Problem 4, $HSD = 6.2849$, $p < 0.01$.

The three-way interaction of group size x arrays x problems was significant $F(9,144) = 2.72$, $p < 0.006$. Means and critical HSDs for this interaction are found in Table C1 of Appendix C.

The four-way interaction of group size x vocalization x arrays x problems was significant $F(9,144) = 2.16$, $p < 0.03$. Means and critical HSDs for this interaction are found in Table C2 of Appendix C.

Transfer session

Data from the transfer session were analyzed by four separate two (individual vs. pair) x two (vocalization vs. nonvocalization) x four (arrays) x four (problems) split-plot ANOVAs, with problems treated as a within-subjects variable. One ANOVA was performed for each dependent measure. Tukey's HSDs were used as the tests of multiple comparisons among means. ANOVA summary tables for each analysis are found in Appendix B. Means for each dependent measure on each problem are found in Table 13, while means for each dependent measure on each array are found in Table 14.

Trials to solution   The between-subjects main effect of arrays was significant $F(3,48) = 3.49$, $p < 0.03$. The cars array required more trials to solution than the forms or borders arrays, $HSD = 0.9129$, $p < 0.05$.

Table 13. Means of all dependent measures for problems in transfer session of Experiment 3

Measures	Problems			
	1	2	3	4
Trials to solution	5.016[a]	4.750[a]	4.391	4.000
Untenable hypotheses	0.840	0.783	0.633	0.527[b]
Strategy efficiency	1.570[c]	1.688	1.704	1.822
Decision time	31.641[d]	24.828	25.141	21.156

[a]Significantly more than Problem 4 at $p < 0.05$.

[b]Significantly less than Problem 1 at $p < 0.01$, and Problem 2 at $p < 0.05$.

[c]Significantly less than Problem 4 at $p < 0.01$.

[d]Significantly more than any other Problem at $p < 0.01$.

Table 14. Means of all dependent measures for arrays in the transfer session of Experiment 3

Measures	Arrays			
	Forms	Cars	Borders	Houses
Trials to solution	4.156	5.109[a]	4.203	4.688
Untenable hypotheses	0.647	0.917	0.760	0.458
Strategy efficiency	1.748	1.583	1.664	1.789
Decision time	20.531	32.922[b]	26.734	22.578

[a]Significantly more than Forms or Borders at $p < 0.05$.

[b]Significantly more than Forms or Houses at $p < 0.01$.

The within-subjects main effect of problems was significant $F(3,144) = 7.71$, $p < 0.0001$. Problem 4 required fewer trials to solution than Problems 1 or 2, HSD = 0.5769, $p < 0.05$.

The two-way interaction of vocalization x problems was significant $F(3,144) = 4.26$, $p < 0.007$. Means for this interaction are found in Table 15. Vocalizing subjects used more trials on Problem 1 than on Problems 3 or 4, and than nonvocalizing subjects used on Problem 4, HSD = 1.1215, $p < 0.01$. Vocalizing subjects also required more trials on Problem 1 than on Problem 2, HSD = 0.9642, $p < 0.05$. Nonvocalizing subjects required more trials on Problem 2 than Problem 4, HSD = 1.1215, $p < 0.01$. Nonvocalizing subjects also required more trials on Problem 2 than vocalizing subjects required on Problems 3 or 4, HSD = 0.9642, $p < 0.05$.

Table 15. Mean trials to solution for vocalization x problems in the transfer session of Experiment 3

Vocalization	Problem			
	1	2	3	4
No	4.719	5.063[a]	4.781	3.906
Yes	5.313[b]	4.438	4.000	4.094

[a]Significantly more than nonvocalizing for Problem 4 at $p < 0.01$, and vocalizing for Problems 3 or 4 at $p < 0.05$.

[b]Significantly more than vocalizing for Problems 3 or 4 and nonvocalizing for Problem 4 at $p < 0.01$, and more than vocalizing for Problem 2 at $p < 0.05$.

The two-way interaction of arrays x problems was also significant $F(9,144) = 2.48$, $p < 0.02$. Means for this interaction are found in Table 16. Problem 1 of the cars array required more trials to solution than Problem 4, than any problem of the forms and borders arrays, and than Problems 1 or 4 of the houses array, HSD = 1.7449, $p < 0.01$. Problem 1 of the cars array also required more trials than Problem 3, and than Problems 2 or 3 of the cars array required more trials than Problem 4, and than Problem 4 of the forms array, HSD = 1.7449, $p < 0.01$.

Table 16. Mean trials to solution for arrays x problems in the transfer session of Experiment 3

Arrays	Problem			
	1	2	3	4
Forms	4.688	4.063	4.125	3.750
Cars	6.438[a]	5.438[b]	4.750	3.813
Borders	4.250	4.625	3.875	4.063
Houses	4.679	4.875	4.813	4.375

[a]Significantly more than Cars 4, Forms 1 through 4, Borders 1 through 4, and Houses 1 or 4 at $p < 0.01$, and more than Cars 3 and Houses 2 or 3, $p < 0.05$.

[b]Significantly more than Cars 4 and Forms 4 at $p < 0.01$.

<u>Untenable hypotheses</u> The between-subjects main effect of vocalization was significant $F(1,48) = 14.96$, $p < 0.0003$, with mean transformed proportions of untenable hypotheses for vocalizing and nonvocalizing of 0.4481 and 0.9428, respectively.

The within-subjects main effect of problems was significant $F(3,144) = 4.59$, $p < 0.005$. Subjects used a lower proportion of untenable hypotheses for Problem 4 than for Problem 1, HSD = 0.2924, $p < 0.01$, or than Problem 2, HSD = 0.2412, $p < 0.05$.

The two-way interaction of group size x arrays was significant $F(3,48) = 2.93$, $p < 0.05$. Means for this interaction are found in Table 17. No significant HSDs were obtained.

Table 17. Mean untenable hypotheses for group size x arrays in the transfer session of Experiment 3

Group Size	Arrays			
	Forms	Cars	Borders	Houses
Individuals	0.836	0.979	0.437	0.392
Pairs	0.459	0.855	1.084	0.523

Strategy efficiency The between-subjects main effect of vocalization was significant $F(1,48) = 4.28$, $p < 0.05$, with mean transformed proportion of efficient selections for vocalizing and nonvocalizing of 1.7777 and 1.6143, respectively.

The within-subjects main effect of problems was significant $F(3,144) = 5.39$, $p < 0.002$. Subjects used a higher strategy efficiency for Problem 4 than Problem 1, HSD = 0.1954, $p < 0.01$.

The two-way interaction of arrays x problems was also significant $F(9,144) = 2.55$, $p < 0.01$. Means for this interaction are found in Table 18. Subjects used a lower strategy efficiency for Problem 1 of of the cars array than Problem 4, than Problems 3 or 4 of the forms array, and than Problem 4 of the houses array, HSD = 0.4876, $p < 0.01$. Subjects used a lower strategy efficiency for Problem 1 of the cars array than for Problem 2 of the borders array, HSD = 0.4307, $p < 0.05$. Finally, subjects used a lower strategy efficiency for Problem 2 of the cars array than for Problem 4 of the houses array, HSD = 0.4307, $p < 0.01$.

Table 18. Mean strategy efficiencies for arrays x problems in the transfer session of Experiment 3

Arrays	Problem			
	1	2	3	4
Forms	1.545	1.720	1.863	1.865
Cars	1.334[a]	1.514[b]	1.579	1.906
Borders	1.643	1.799	1.654	1.561
Houses	1.758	1.719	1.721	1.956

[a]Significantly less than Cars 4, Forms 3 or 4, and Houses 4 at $p < 0.01$, and less than Borders 2 at $p < 0.05$.

[b]Significantly less than Houses 4 at $p < 0.01$.

Decision time The between-subjects main effects of arrays was significant $F(3,48) = 6.73$, $p < 0.0007$. The cars array elicited more decision time than the forms or houses arrays, HSD = 9.9063, $p < 0.01$.

The within-subjects main effect of problems was significant $F(3,144) = 10.44$, $p < 0.0001$. Problem 1 required more decision time than any other, HSD = 5.935, $p < 0.01$.

Transfer across sessions

The analysis of data reflecting changes in performance from the learning to transfer sessions consisted of four separate two (individual vs. pair) x two (vocalization vs. nonvocalization) x two (grouped order: semantic-to-perceptual vs. perceptual-to-semantic) x two (learning vs. transfer session) split-plot ANOVAs, with sessions treated as a within-subjects variable. One ANOVA was completed for each dependent measure. Tukey's HSDs were used to test multiple comparisons of means. Summary tables of all ANOVAs are found in Appendix B.

Trials to solution The between-subjects main effect of group size was significant $F(1,56) = 4.61$, $p < 0.04$, with mean trials to solution for individuals and pairs of 19.7188 and 17.7344, respectively.

The two-way interaction of vocalization x grouped arrays was significant $F(1,56) = 7.60$, $p < 0.008$. Means for this interaction are found in Table 19. Nonvocalizing subjects in the semantic-to-perceptual condition required more trials than vocalizing subjects in the same condition, HSD = 3.5015, $p < 0.05$.

Table 19. Mean trials to solution for vocalization x grouped arrays

Grouped Arrays	Vocalization	
	No	Yes
Perceptual-to-semantic	17.813	19.094
Semantic-to-perceptual	20.906[a]	17.094

[a]Significantly more than semantic-to-perceptual/vocalization at $p < 0.05$.

The two-way interaction of group size x session was also significant $F(1,56) = 4.40$, $p < 0.05$. Means for this interaction are depicted in Figure 7. Individually trained subjects required more trials during the learning session than any other condition, HSD = 2.9095, $p < 0.05$.

The two-way interaction of grouped arrays x session was also significant $F(1,56) = 17.07$, $p < 0.0001$. Means for this interaction are found in Table 20. Subjects in the learning session of the semantic-to-perceptual condition required more trials than in the transfer session, and than subjects in the learning session of the perceptual-to-semantic condition, HSD = 3.6082, $p < 0.01$.

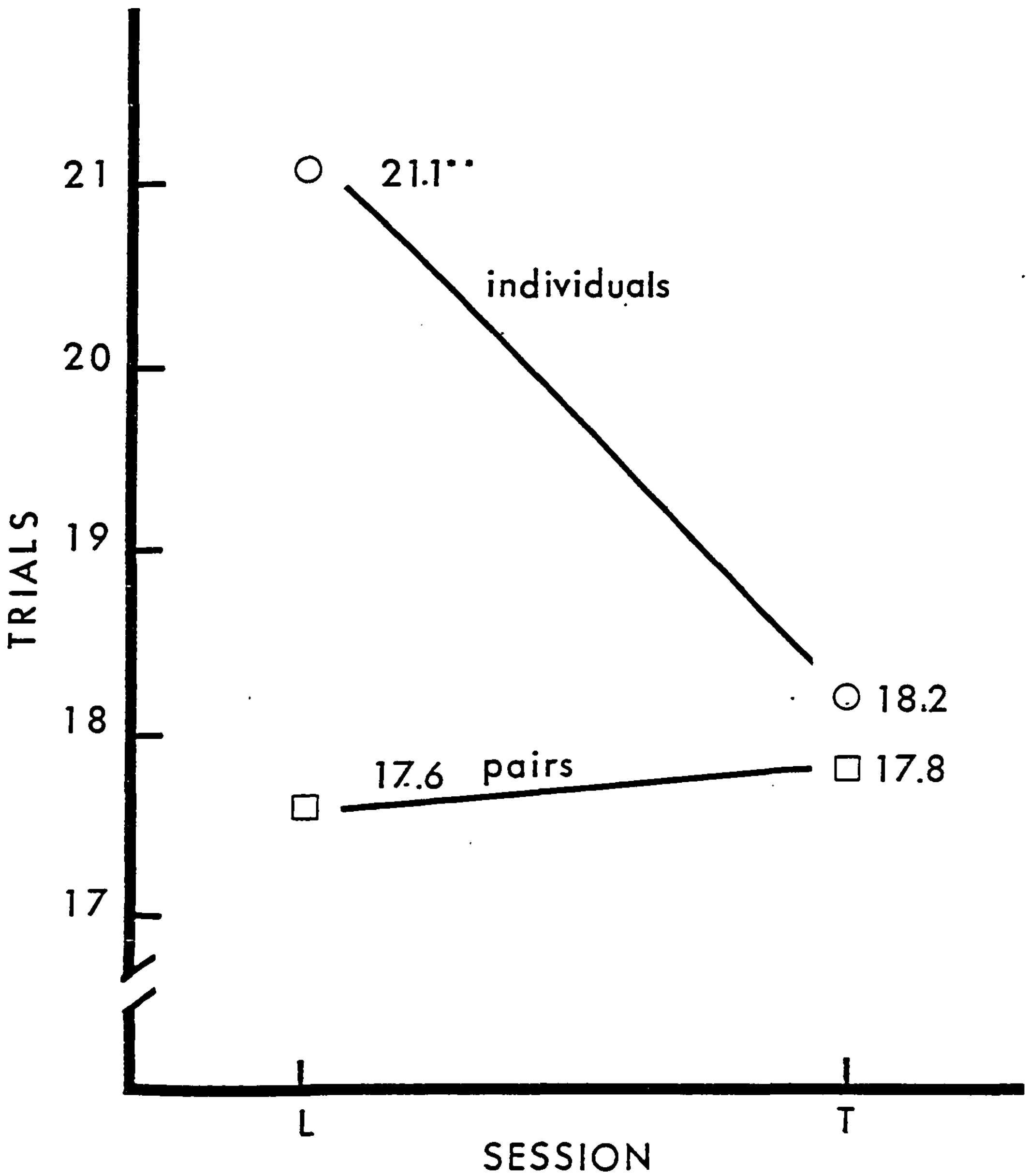
Table 20. Mean trials to solution for grouped arrays x sessions

Grouped Arrays	Sessions	
	Learning	Transfer
Perceptual-to-semantic	17.563	19.344
Semantic-to-perceptual	21.281[a]	16.719

[a]Significantly more than semantic-to-perceptual/transfer and perceptual-to-semantic/learning at $p < 0.01$.

Untenable hypotheses The between-subjects main effect of vocalization was significant $F(1,56) = 12.62$, $p < 0.0008$, with mean transformed proportions of untenable hypotheses for vocalizing and non-vocalizing of 0.5653 and 0.9689, respectively.

The within-subjects main effect of sessions was significant $F(1,56) = 7.97$, $p < 0.007$, with means for the learning and transfer sessions of 0.6776 and 0.8566, respectively.



**more than any other at 0.01

FIGURE 7. Mean Trials to Solution for Group Size x Session in Experiment 3

The two-way interaction of group size x sessions was significant $F(1,56) = 7.54$, $p < 0.009$. Means for this interaction appear in Figure 8. Subjects trained in pairs used a lower proportion of untenable hypotheses in the learning session than in the transfer session, and than individually trained subjects used in either session, HSD = 0.2981, $p < 0.01$.

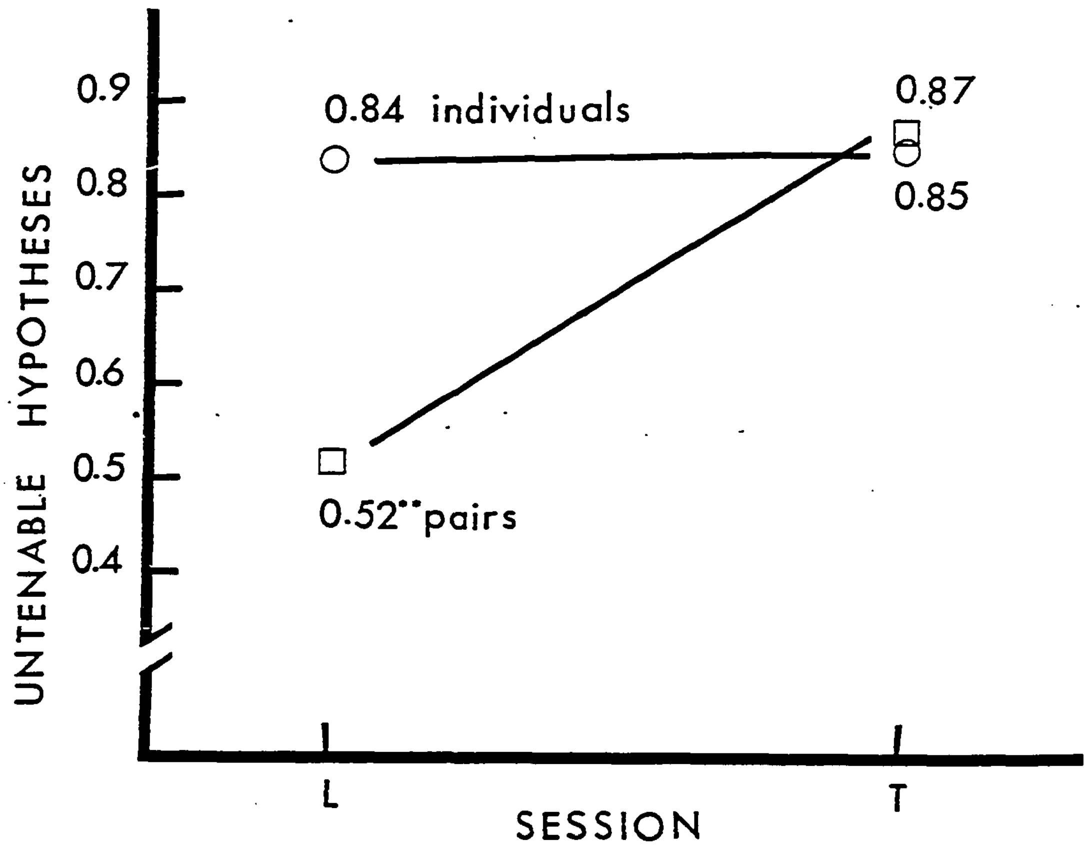
The two-way interaction of grouped arrays x sessions was also significant $F(1,56) = 24.48$, $p < 0.001$. Means for this interaction are found in Table 21. Subjects in this learning session of the perceptual-to-semantic condition used a lower proportion of untenable hypotheses than in the transfer session, and than subjects in the semantic-to-perceptual condition in either session, HSD = 0.2981, $p < 0.01$.

Table 21. Mean untenable hypotheses for grouped arrays x sessions

Grouped Arrays	Sessions	
	Learning	Transfer
Perceptual-to-semantic	0.422[a]	0.914
Semantic-to-perceptual	0.934	0.799

[a]Significantly less than any other mean at $p < 0.01$.

Strategy efficiency The between-subjects main effect of group size was significant $F(1,56) = 4.07$, $p < 0.05$, with mean transformed proportions of efficient selections for individuals and pairs of 1.6203 and 1.7453, respectively. The between-subjects main effect of vocalization was also significant $F(1,56) = 6.37$, $p < 0.02$, with mean strategy



** less than any other at 001

FIGURE 8. Mean Untenable Hypotheses for Group Size x Session in Experiment 3

efficiencies for vocalizing and nonvocalizing of 1.7609 and 1.6046, respectively.

The within-subjects main effect of sessions was significant $F(1,56) = 4.87$, $p < 0.04$, with mean strategy efficiencies for the learning and transfer sessions of 1.7180 and 1.6476, respectively.

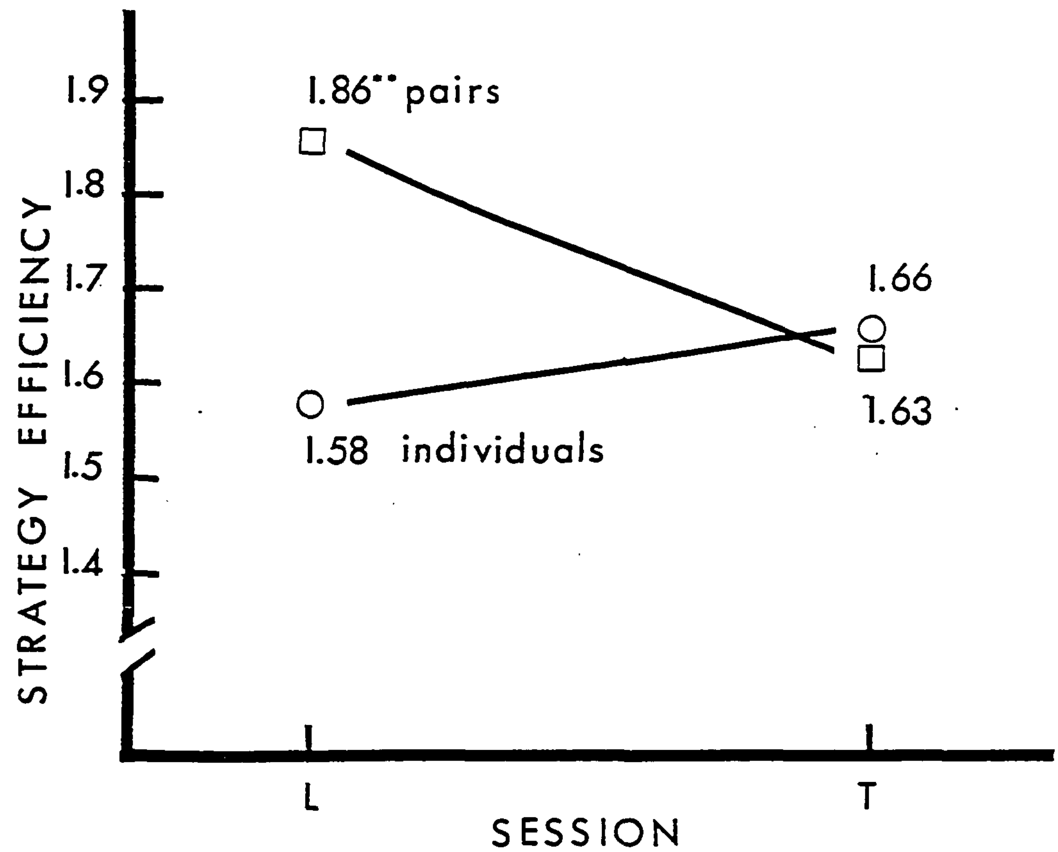
The two-way interaction of group size x sessions was significant $F(1,56) = 23,89$, $p < 0.0001$. Means for this interaction are found in Figure 9. Subjects trained in pairs used a greater proportion of efficient selections during the learning session than during the transfer session. Pairs, in the learning session, also used a greater strategy efficiency than individuals in either session, HSD = 0.1498, $p < 0.01$.

The two-way interaction of grouped arrays x session was also significant $F(1,56) = 30.64$, $p < 0.0001$. Means for this interaction are found in Table 22. Subjects in the learning session of the perceptual-to-semantic condition used a greater proportion of efficient selections than in the transfer session, and than subjects in the semantic-to-perceptual condition in either session, HSD = 0.1498, $p < 0.01$.

Table 22. Mean strategy efficiencies for grouped arrays x sessions

Grouped Arrays	Sessions	
	Learning	Transfer
Perceptual-to-semantic	1.866[a]	1.619
Semantic-to-perceptual	1.570	1.676

[a]Significantly more than any other mean at $p < 0.01$.



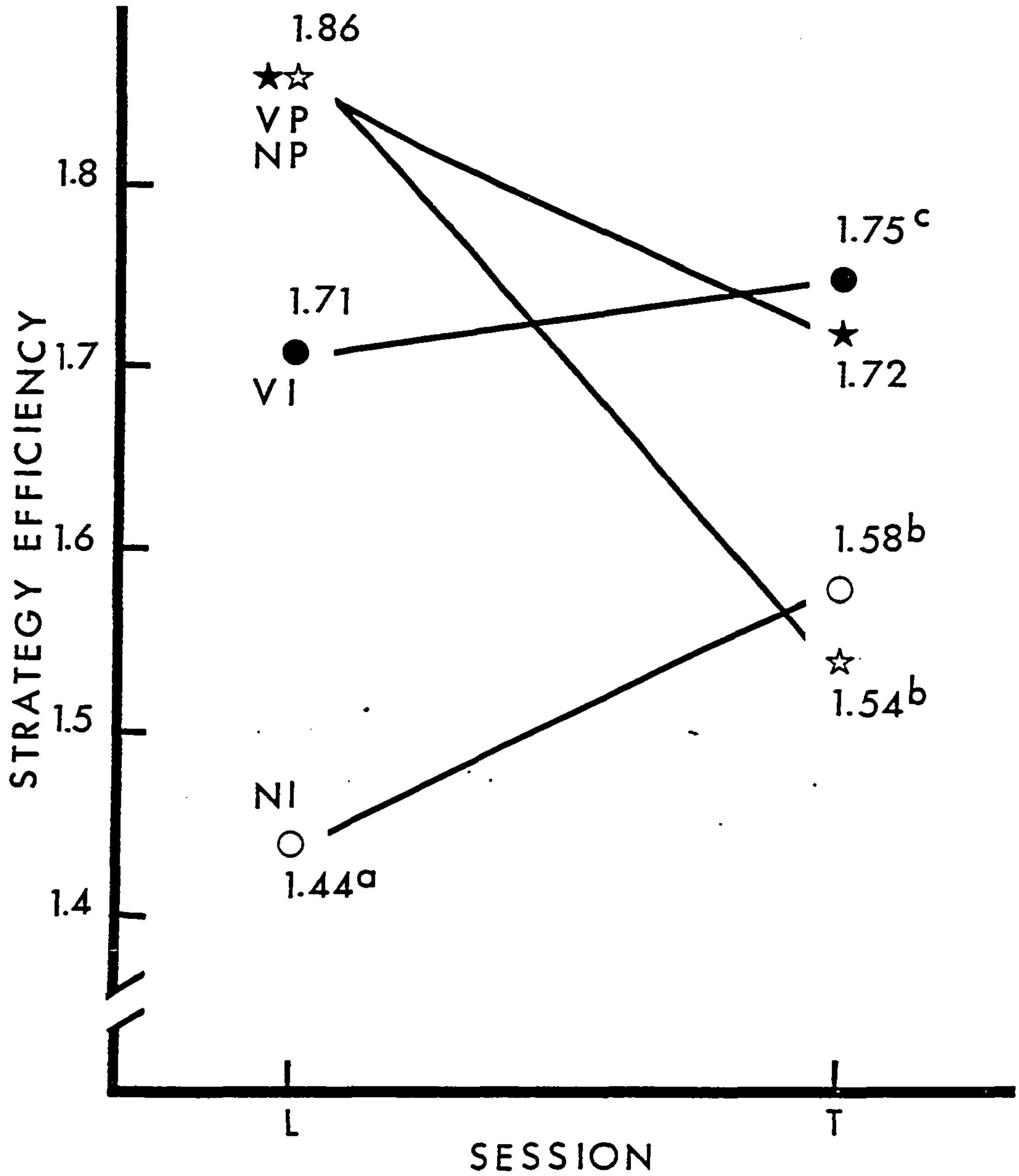
** more than any other at 0.01

FIGURE 9. Mean Strategy Efficiencies for Group Size x Session in Experiment 3

The three-way interaction of group size x vocalization x session was also significant $F(1,56) = 4.78$, $p < 0.04$. Means for this interaction are depicted in Figure 10. Nonvocalizing individuals, during the learning session, used a lower proportion of efficient selections than all other conditions, except nonvocalizing pairs and individuals, during the transfer task, HSD = 0.2429, $p < 0.01$. Subjects trained as nonvocalizing pairs used less strategy efficiency, in the transfer session, than nonvocalizing and vocalizing pairs used in the learning session, HSD = 0.2429, $p < 0.01$. Subjects trained as nonvocalizing pairs also used less strategy efficiency, in the transfer session, than did individually trained vocalizing subjects, HSD = 0.2037, $p < 0.05$. Individually trained nonvocalizing subjects used less strategy efficiency, in the transfer session, than did vocalizing and nonvocalizing pairs, during the learning session, HSD = 0.2429, $p < 0.01$.

Decision time The between-subjects main effect of group size was significant $F(1,56) = 6.00$, $p < 0.02$, with mean decision times for individuals and pairs of 29.9631 and 35.0313 seconds, respectively. The between-subjects main effect of vocalization was also significant $F(1,56) = 10.75$, $p < 0.02$, with mean decision times for vocalizing and nonvocalizing of 35.8906 and 29.0938 seconds, respectively.

The within-subjects main effect of session was also significant $F(1,56) = 100.40$, $p < 0.0001$, with mean decision times for the learning and transfer sessions of 39.0156 and 25.9688 seconds, respectively.



[a] less than any vocalizing group at 0.01
[b] less than VP or NP at L at 0.01
[c] more than NP at T at 0.05

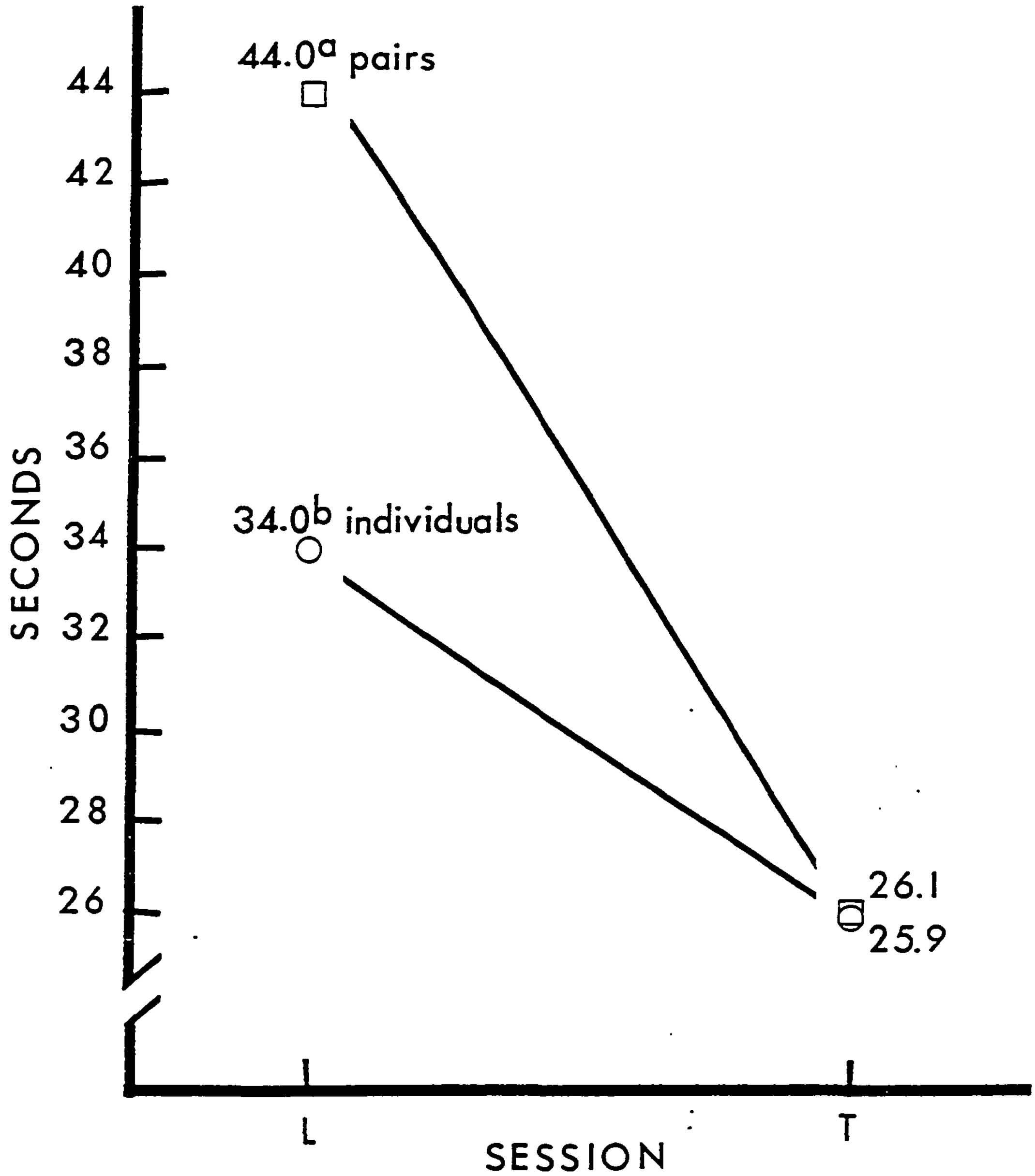
Note. V = vocal, N = non-vocal, I = individuals, P = pairs.

FIGURE 10. Mean Strategy Efficiencies for Group Size x Vocalization x Session in Experiment 3

The two-way interaction of group size x session was significant $F(1,56) = 14.11$, $p < 0.0004$. Means for this interaction are found in Figure 11. Subjects trained in pairs used more decision time in the learning session than in the tranfer session, and than individually trained subjects in either session, HSD = 6.1199, $p < 0.01$. Individually trained subjects used more decision time in the learning session than either group size in the transfer session, HSD = 6.1199, $p < 0.01$.

The two-way interaction of vocalization x session was significant $F(1,56) = 13.57$, $p < 0.0005$. Means for this interaction are found in Figure 12. During the learning session, subjects who vocalized required more decision time than any group, HSD = 6.1199, $p < 0.01$. Nonvocalizing subjects required more decision time in the learning sesssion than either group in the transfer session, HSD = 6.1199, $p < 0.01$.

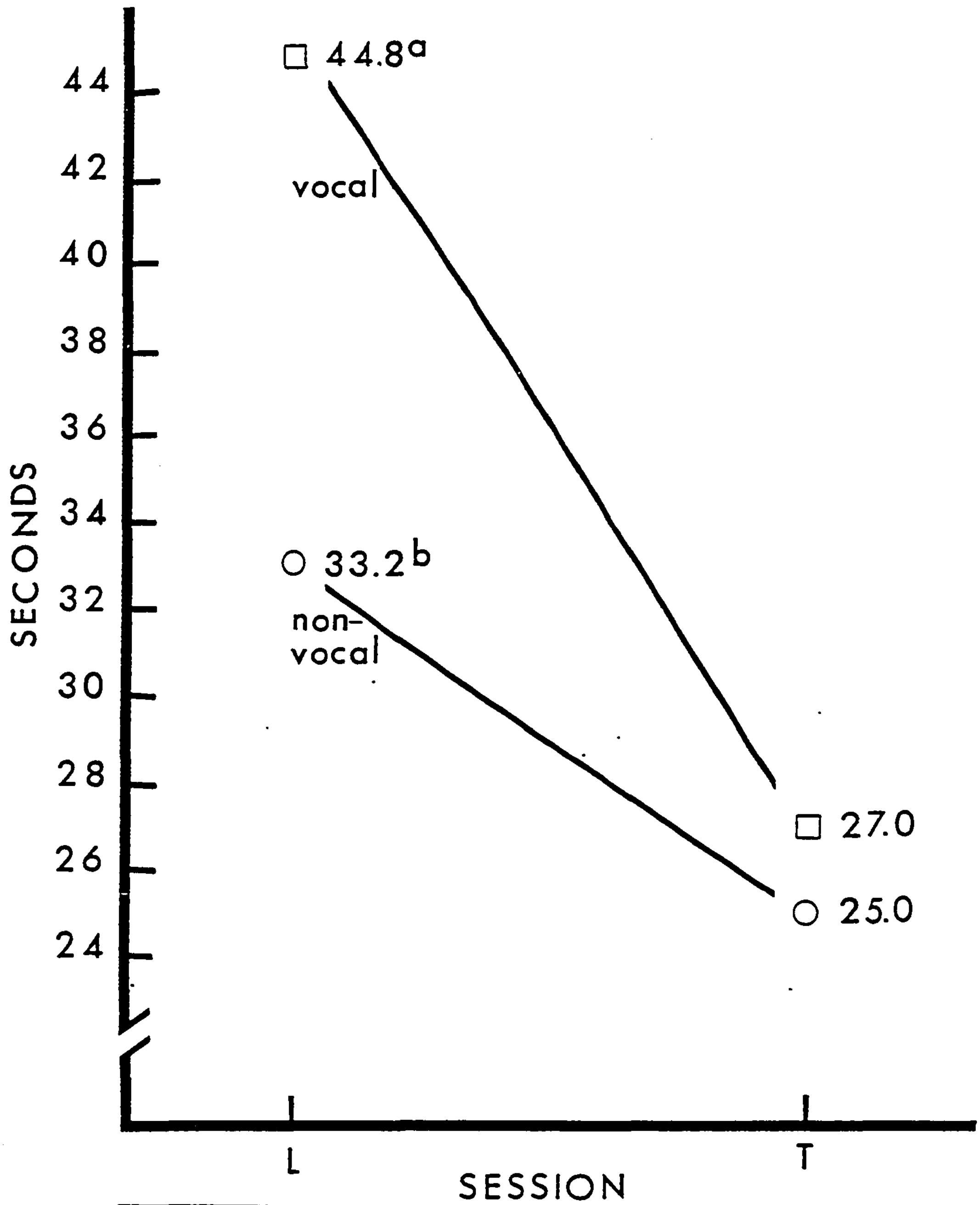
The two-way interaction of grouped arrays x session was also significant $F(1,56) = 12.70$, $p < 0.0008$. Means for this interaction are found in Table 23. Subjects in the learning session of both conditions required more decision time than in the transfer session, HSD = 6.1199, $p < 0.01$. Also, subjects in the transfer session of the perceptual-to-semantic condition required more decision time than subjects from the semantic-to-perceptual condition in the same session, HSD = 4.935, $p < 0.05$.



[a] more than any other at 0.01

[b] more than any T at 0.01

FIGURE 11. Mean Decision Times for Group Size x Session in Experiment 3



[a] more than any other at 0.01
[b] more than any T at 0.01

FIGURE 12. Mean Decision Times for Vocalization x Session in Experiment 3

Table 23. Mean decision times for grouped arrays x sessions

Grouped Arrays	Session	
	Learning	Transfer
Perceptual-to-semantic	36.938[a]	28.531[b]
Semantic-to-perceptual	39.656[a]	23.406

[a]Significantly more than any transfer score at $p < 0.01$.

[b]Significantly more than semantic-to-perceptual/transfer at $p < 0.05$.

The three-way interaction of group size x grouped arrays x session was also significant $F(1,56) = 4.41$, $p < 0.05$. Means for this interaction are found in Table C3 of Appendix C. Paired subjects in the learning session of the semantic-to-perceptual condition required more decision time than any previously paired subjects in the transfer session, and than any individually trained group in either session, HSD = 9,9255, $p < 0.01$. Paired subjects in the learning session of the perceptual-to-semantic condition required more decision time than any group in the transfer session, HSD = 9.9255, $p < 0.01$. Individual subjects in the learning session of the semantic-to-perceptual condition required more decision time than subjects trained as individuals or pairs required in the transfer session of the same condition, HSD = 9.9255, $p < 0.01$. Individual subjects in the learning session of the perceptual-to-semantic condition required more decision time than did previously paired subjects in the transfer session of the semantic-to-perceptual condition, HSD = 9.9255, $p < 0.01$; or than individually trained subjects in the transfer session of the semantic-to-perceptual condition, HSD = 8.3234, $p < 0.05$.

Discussion

Solving a series of rule-learning problems did not improve the performance on subsequent problems in all cases. Although performance within each session generally improved from problem to problem, the transfer from session to session was dependent on group size and vocalization. In general, subjects trained in pairs evinced negative transfer of rule-learning processing between sessions. Vocalizing the reasoning during the initial session improved the processing in both sessions, with the effect of decreasing the severity of the negative transfer for previously paired subjects.

Free discussion with a partner did not affect rule-learning in the same manner as vocalizing the reasoning. In the learning session, working with a partner elicited fewer trials to solution, but more decision time than rule-learning as individuals. Also, nonvocalizing pairs were better at interpreting feedback and testing hypotheses than nonvocalizing individuals. However, the group size differences did not appear in the transfer session. In the transfer session, vocalization was the major factor affecting results. Subjects who vocalized in the learning session were better at interpreting feedback and testing hypotheses than subjects who did not vocalize their reasoning. Group size had some effect, however, as subjects trained as nonvocalizing pairs used less strategy efficiency in the transfer sesion than did subjects trained as vocalizing individuals.

Metacognitively, working with a partner is analogous to monitoring by another, while vocalization requires a degree of self-monitoring of reasoning. Thus, monitoring by another improved the performance of the group, but not the subsequent performance of individual members of the group. Self-monitoring through vocalization improved the rule-learning processing both while the subjects vocalized and during a subsequent task involving no overt monitoring.

Transfer on successive problems

Positive near transfer was reflected by trials to solution, strategy efficiency, and decision time in the learning session. Positive near transfer within the transfer session was reflected by all four dependent measures.

Learning session   More trials to solution were required for Problem 1 of the cars array than for any other problem of any array. Vocalizing and paired subjects required more decision time for Problem 1 than other problems, and than individuals and nonvocalizing subjects required for most problems. Poorer strategy efficiency was used on Problem 1 than on Problems 3 or 4.

Transfer session   Vocalizing subjects required more trials to solution for Problem 1 than Problems 2 through 4, or than nonvocalizing subjects required for Problem 4. Nonvocalizing subjects required more trials for Problem 2 than Problem 4, or than vocalizing subjects required for Problems 3 or 4. Problem 1 of the cars array required more trials than Problems 3 or 4, or any problems from other arrays. Problem 2 of the cars array required more trials than Problem 4 of the cars or forms arrays. Poorer strategy efficiency was used for Problem 1 of the cars array than for Problem 4 of the cars array, Problem 2 of the borders array, Problems 3 or 4 of the forms array, and Problem 4 of the houses array. Poorer strategy efficiency was also used for Problem 2 of the cars array than for Problem 4 of the houses array. Subjects used a greater proportion of untenable hypotheses for Problem 1 than Problem 4. Problem 1 required more decision time than any other problem.

Transfer across sessions

The overall transfer between the learning and transfer sessions, across groups size, vocalization, and presentation order, was not positive. The transfer session required more decision time, a greater proportion of untenable hypotheses and less strategy efficiency than the learning session. However, the sessions main effects were invariably part of some interaction.

Array difficulty and presentation order The most difficult array in either session was the semantically-oriented cars array. In the learning session, subjects used a higher proportion of untenable hypotheses for the cars array than for any other array. The cars and houses arrays required less strategy efficiency than the borders array. In the transfer session, the cars array required more decision time than did the forms or houses arrays.

Greater difficulty of the semantic arrays in general was also evident. In the learning session, the perceptual arrays resulted in fewer trials to solution, better processing of feedback and higher strategy efficiency than did the semantic arrays. The semantic-to-perceptual presentation order resulted in fewer trials to solution and less decision time in the transfer session than in the learning session. The perceptual-to-semantic presentation order resulted in poorer rule-learning in the transfer session than in the learning session.

Group size   Rule-learning with a partner led to generally poorer processing in the transfer session than in the learning session. Previously paired subjects evinced a decrement in the accuracy of interpreting feedback and the efficiency of hypothesis-testing, when tested as individuals. Individually trained subjects evinced no change in processing, but used less decision time and fewer trials to solution in the transfer session than in the learning session. Previously paired subjects also used less decision time in the transfer session than in the learning session.

Differences between pairs and individuals in the learning session disappeared in the transfer session. Initial rule-learning with a partner resulted in more decision time, better processing and fewer trials to solution than initial rule-learning as individuals. However, the superiority of the pairs did not transfer to the individual transfer session.

Vocalization   Vocalizing the reasoning affected performance in both sessions. Although the pattern of decision time results for vocalizing reasoning during rule-learning was similar to that of working with a partner, the effect of vocalizing on rule-learning processing was generally positive. In the learning session, nonvocalizing individuals were poorer at interpreting feedback and testing hypotheses than any other group, including nonvocalizing pairs. However, in the transfer session, subjects from both nonvocalizing groups used poorer processing than previously vocalizing groups. Although the positive influence of working with a partner was apparent only during the learning session, the positive effect of vocalizing the reasoning during the learning session transferred to the nonvocalized transfer session.

The three-way interaction of vocalization x group size x sessions, reflected by strategy efficiency, illustrated the differential effects of vocalizing and working with a partner. In the learning session, individuals who did not vocalize used the poorest strategy efficiency. In the transfer session, a condition in which working with a partner had a negative effect, the performance of previously paired subjects declined. However, previously paired vocalizing subjects evinced a less severe decline than previously paired nonvocalizing subjects.

Conclusion

Positive near transfer occurred in both sessions. However, the overall transfer from session to session depended on group size, vocalization, and the presentation order of the arrays. In the learning session, rule-learning as individuals required more trials to solution than rule-learning with a partner. The rule-learning processing of individuals, especially nonvocalizing individuals, was also poorer than that of pairs. This superiority of the group processing may have been due to monitoring by another or self-monitoring. Both vocalizing and working with a partner required more decision time, presumably because the acts of stating the reasoning and discussion required more time.

In the transfer session, individually trained subjects decreased their number of trials to solution to the same level as previously paired subjects. Meanwhile, the rule-learning processing of previously paired subjects also decreased to a level similar to that of individually trained subjects. The general effect of vocalization, however, was to keep the processing of all subjects at a higher level than that of nonvocalizing subjects. With active vocalization and the discussion of the group removed, the decision time of all subjects in the transfer session was less than that in the learning session.

The difficulty of the stimuli had little effect on the above results, however, rule-learning was affected by the type of stimulus array. Perceptually-oriented arrays were more easily processed than the more semantically-oriented arrays; and required fewer trials to solution, when used in the learning session. The difficulty of the semantic arrays may have been due to the greater difficulty of the cars array.

In summary, the effect of working with a partner was better initial general performance, but a subsequent decline in individual processing.

The effect of vocalizing the reasoning in the learning session was to improve the individual processing of subjects trained as individuals or in pairs, during both sessions. These results occurred over perceptually- and semantically-oriented stimuli of different levels of difficulty.

GENERAL DISCUSSION

Transfer was evident both between problems and sessions. Positive transfer between problems was reflected most often by decision time, followed by strategy efficiency and trials to solution. The proportion of untenable hypotheses rarely reflected positive transfer from problem to problem. Individuals generally evinced positive transfer between sessions. However, transfer for paired subjects indicated a drop in processing from the learning to the transfer session. When the paired subjects were required to vocalize their reasoning in the learning session, the decline in processing was less severe. Vocalization of reasoning improved the performance of individually trained and paired subjects in both sessions. Although both vocalization and working with a partner improved performance, these conditions also required more decision time than nonvocalization and individual rule-learning, respectively. No effects for the sex of the subject were found. These results occurred over a variety of stimulus material.

Difficulty of the Arrays

No differential transfer between the learning and transfer sessions occurred for the presentation order of the grouped arrays, in Experiment 2. However, the more powerful Experiment 3 resulted in some differential transfer. In general, the semantic arrays were more difficult, especially the cars array. This difficulty of the semantic arrays was most pronounced in the learning session. Apparently, the difficulty of the arrays had no effect on the cognitive monitoring of rule-learning other than the more difficult arrays eliciting poorer performance.

Group Size

Effects of group size were most prevalent in the learning session. Pairs required more decision time than individuals. Individually trained subjects required more trials to solution than did pairs, across a variety of arrays. Individuals, especially nonvocalizing individuals, also performed worse than pairs at testing hypotheses and interpreting feedback, across a variety of arrays. Thus, the general effect of working with a partner appeared to be improved processing and overall performance, but increased decision time required for processing.

The only significant effects involving group size in the transfer session occurred in Experiment 1. Individually trained subjects were better at testing hypotheses and interpreting feedback than previously paired subjects. Experiment 1 differed from Experiment 3 in that the former used only the forms to cars presentation order. Apparently, the forms array, when presented alone, resulted in no differences in performance between individuals and pairs. Therefore, the decrement in performance in the transfer session for previously paired subjects resulted in inferior scores for the proportion of untenable hypotheses and strategy efficiency.

Differential transfer between sessions was evident for pairs and individuals. In Experiment 3, the inferior general performance of individuals and the superior processing of pairs, present across arrays in the learning session, disappeared in the transfer session. Trials to solution reflected improvement for individually trained subjects. Processing latency improved for both individually trained subjects and those trained in pairs. Subjects trained in pairs evinced a decline in the efficiency of hypothesis-testing and the interpretation of feedback.

Most of the effects of group size which favored rule-learning with a partner occurred only as long as the group was together. When members of the pair were separated, processing declined. Experiment 1 illustrated that, under some circumstances, this decrement can be severe enough for individually trained subjects to evince more effective processing than previously paired subjects. Monitoring by another did not improve individual cognitive monitoring, as reflected in the transfer between sessions.

Vocalization

The self-monitoring induced by vocalization of reasoning improved the overall performance in both sessions. However, actively vocalizing subjects required more decision time than nonvocalizing subjects. The rule-learning processing of actively vocalizing subjects was superior to that of nonvocalizing individuals; while there was no difference between the processing of vocalizing and nonvocalizing pairs, in the learning session. In the transfer session, previously vocalizing subjects used more effective processing than did nonvocalizing subjects, across group size.

The between-sessions transfer effects of vocalization differed from those of group size. In the learning session, either vocalizing or working with a partner improved the cognitive monitoring of paired subjects. If working with a partner was the condition which improved the personal cognitive monitoring of each member of the pair, there should have been little decrement in the processing in the individual transfer task. Results of group size indicated a decrement in processing for subjects trained in pairs. However, previously vocalizing pairs did not evince as sharp a decline in efficient hypothesis-testing as did non-vocalizing subjects. Apparently, vocalizing with a partner decreased the negative effect rule-learning in pairs had on subsequent individual processing.

Problems

Results of this study may have been affected by "lucky guesses." A lucky guess is a correct conclusion offered before it could be logically considered the only correct choice. Lucky guesses may be used to reduce the number of trials to solution, to gain more information per trial, or because the learner sincerely believes they are the only possible rules. Whatever the reason, the lucky guess does affect the dependent measures. The number of trials to solution is reduced with the lucky guess. If it is assumed that the lucky guess would have been used as a possible rule in a normal trial, the lucky guess would have two possible effects on the proportion of untenable hypotheses. If no untenable hypotheses have been made before the lucky guess, there would be no change in score. If at least one untenable hypothesis has occurred, the use of the lucky guess increases the proportion of untenable hypotheses. The use of the lucky guess may increase or decrease the strategy efficiency score, depending on whether the selected instance which would have been tested was informative or not. Methods for curbing the effects of the lucky guess have been offered (Johnson, 1978). However, these procedures remove the laboratory situation even further from more natural learning activities, in which lucky guesses may occur.

Another problem concerns the relationship of the proportion of untenable hypotheses to strategy efficiency. The existence of untenable hypotheses is the limiting factor of the strategy efficiency score. However, it is possible that the testing of hypotheses and the interpretation of feedback are not as logically connected in subjects' processing as they are in the logic of the scoring procedures. A measure of strategy efficiency which is independent of the proportion of untenable hypotheses may be desirable. If such a measure contained a provision for some scoring penalty for lucky guesses, subjects' natural processing would be allowed, yet the score would account for lucky guesses.

Conclusion

Results of this study indicate that one aspect of working with a partner which leads to superior small group performance may be the reciprocal monitoring of the partners' processing. However, monitoring by others does not prevent a decline in processing when members of the small group are subsequently tested as individuals. Self-monitoring, as induced by vocalization of reasoning, may reduce the decline in performance evinced by previously paired subjects. In other words, individuals may initially perform poorer than small groups because of a lack of monitoring by others. However, individuals must rely on self-monitoring, and thus, do not show the same decline in performance as previously paired subjects, when tested individually on a similar task.

These results have implications for learning research. Studies of the effects of small group learning should include a transfer session. This study demonstrated that superior performance of paired subjects does not necessarily transfer to individual performance. This finding would not have been discovered without the use of a transfer task. Furthermore, if learners are allowed to work on a task in small groups, some procedure, such as vocalization of reasoning, which fosters the self-monitoring of individual processing should also be included. Without such a procedure, the subsequent processing by individuals from the groups may be less effective than possible.

A distinction can be made between solving a series of problems and applying the solution processes to subsequent problems. The former occurred in the learning session of these studies, while the latter was reflected by the transfer session. In general, subjects trained in pairs did not apply their previously used processing as well as did individuals. It is assumed that self-monitoring, which occurred in the learning session, was responsible for the transfer of problem-solving processes for individually trained subjects. For a variety of reasons, pairs may not have engaged in as much self-monitoring. Social dominance or ability differences may not have allowed the lesser member to develop the necessary processing. Paired subjects may have relied on the resources of others or engaged in social loafing.

Results for vocalization indicate that self-monitoring is one metacognitive variable necessary for the positive transfer of problem-solving processing. Self-monitoring need not take the form of the vocalization of reasoning. Other methods of increasing self-monitoring may be to increase the motivation to perform well, improve the interaction of the group, or directly train subjects in self-monitoring techniques. Self-monitoring should occur whenever a learner is actively participating in the learning process. Thus, individuals who did not vocalize were actively engaged in their learning, because there was no one else to solve the problem for them. Vocalizing individuals were more actively engaged in learning, because vocalizing requires close self-monitoring. Apparently, metacognitive monitoring is not an all-or-none process.

Future areas of investigation into self-monitoring include the nature of self-monitoring, varieties of inducements to self-monitor, and the effectiveness of self-monitoring on a variety of tasks. First, if self-monitoring is interfered with by various social influences, then manipulating those influences should affect subjects' individual performance. For example, subjects from nondiscussing pairs who take turns on each problem and use the information collected by the other should evince less negative transfer than discussing pairs, who did not vocalize. More self-monitoring would be required of the nondiscussing pairs. Second, the type of vocalization may affect performance. Differences in subsequent nonvocalized performance may occur between groups allowed free vocalization, guided vocalization, and those trained in self-monitoring. Finally, a variety of tasks and group sizes should be used in future investigations of self-monitoring.

Results of these experiments also have some implications for the classroom use of groups and self-monitoring. As with laboratory studies, a distinction can be made between problem-solving performance and the ability to retrieve and use previously learned problem-solving processes. Both types of learning are important in the classroom. It has been established that learning in small groups does improve the performance of the group, develop positive social interactions between students, and improve attitudes of students towards school. However, students from small groups may not use previous problem-solving processing as well as do individually trained students. Thus, effective teaching for retrieval of problem-solving processes may entail requiring grouped students to self-monitor, as do individually trained students. Again, this self-monitoring need not take the form of vocalization of reasoning. One method of improving self-monitoring may be to use teams of students, in which each member of the team is responsible for teaching part of the lesson and for sharing how they learned the information.

Extraneous social influences are a greater problem in classroom studies of small group problem-solving than in laboratory studies. Control of social influences in the classroom is probably impossible.

Dominance patterns, sex-related variables, interaction patterns, decision schemes, social climate, motivation, ability differences, and previous learning experiences may all affect small group performance and the self-monitoring of members of the group. This does not mean that self-monitoring in the classroom will not occur; it means that self-monitoring should not be considered the only factor influencing far transfer of problem-solving processes. Future areas of classroom research on self-monitoring and group problem-solving include the above factors, as well as various problem-solving tasks, grade level, and ability groupings.

REFERENCES

Archer, E. J., Bourne, L. E., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. Journal of Experimental Psychology, 1955, 49, 153-164.

Beane, W. E. & Lemke, E. A. Group variables influencing the transfer of conceptual behavior. Journal of Educational Psychology, 1971, 62, 215-218.

Bender, T. A. Dyadic instruction and the transfer of rule-learning strategies. Unpublished Master's Thesis. Iowa State University, Ames, Iowa, 1980.

Bourne, L. E. Human Conceptual Behavior. Boston: Allyn & Bacon, 1966.

Braley, L. S. Strategy selection and negative instances in concept learning. Journal of Educational Psychology, 1963, 54, 154-159.

Brown, A. L. The development of memory: Knowing, knowing about knowing, and knowing how to know. Advances in Child Development and Behavior, 1975, 10, 57-101.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in Instructional Psychology (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A Study of Thinking. New York: Wiley & Sons, 1956.

Byers, J. L. Strategies and learning set in concept attainment. Psychological Reports, 1963, 12, 623-634.

Cahill, H. E. & Hovland, C. I. The role of memory in the acquisition of concepts. Journal of Experimental Psychology, 1960, 59, 137-144.

Cohen, J. & Cohen, P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates, 1975.

Conant, M. B. & Trabasso, T. Conjunctive and disjunctive concept formation under equal-information conditions. Journal of Experimental Psychology, 1964, 67, 250-255.

Davis, J. H., Carey, M. H., Foxman, P. N., & Tarr, D. B. Verbalization, experimenter presence, and problem solving. Journal of Personality and Social Psychology, 1968, 8, 299-302.

Denny, J. P. Effects of anxiety and intelligence on concept formation. Journal of Experimental Psychology, 1966, 72, 596-602.

Dominowski, R. L. Role of memory in concept learning. Psychological Bulletin, 1965, 63, 271-289.

Durling, R. & Schick, C. Concept attainment in pairs and individuals as a function of vocalization. Journal of Educational Psychology, 1976, 68, 83,-91.

Eifermann, R. R. Response patterns and strategies in the dynamics of concept attainment behavior. British Journal of Psychology, 1965, 56, 217-222.

Flavell, J. H. Metacognitive aspects of problem-solving. In L. B. Resnick (Ed.), The Nature of Intelligence, Hillsdale, NJ: Lawrence Erlbaum Associates, 1976.

Flavell, J. H. Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), Structural/Process Theories of Complex Human Behavior. Alphen aan der Rijn, The Netherlands: Sijthoff & Noordhoff, 1978.

Flavell, J. H. Metacognition and cognitive monitoring: A new area of of cognitive-developmental inquiry. American Psychologist, 1979, 34, 906-611.

Freibergs, V. & Tulving, E. The effect of practice on utilization of information from positive and negative instances in concept identification. Canadian Journal of Psychology, 1961, 5, 101-106.

Gagne, R. M. & Smith, E. C. A study of the effects of verbalization on problem-solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Glass, A. L., Holyoak, K., & Santa, J. L. Cognition. Reading, Massachusetts: Addison-Wesley, 1979.

Hagen, J. W., Jongeward, R. H., & Kail, R. V. Cognitive perspectives on the development of memory. Advances in Child Development and Behavior, 1975, 10, 57-101.

Haygood, R. C. & Bourne, L. E. Attribute- and rule-learning aspects of conceptual behavior. Psychological Review, 1965, 72, 175-195.

Hill, G. W. Group versus individual performance: Are N+1 heads better than one? Psychological Bulletin, 1982, 91, 517-539.

Hovland, C. I. & Weiss, W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-182.

Hunt, E. B. Memory effects in concept learning. Journal of Experimental Psychology, 1961, 62, 598-604.

Hunt, E. B. & Hovland, C. I. Order of consideration of different concept types. Journal of Experimental Psychology, 1969, 59, 220-225.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. Psychological Bulletin, 1981, 89, 47-62.

Johnson, E. S. Validation of concept-learning strategies. Journal of Experimental Psychology: General, 1978, 107, 237-266.

Johnson, E. S. Assessment of learning in the selection concept-learning task. Journal of Educational Psychology, 1980, 72, 821-829,

Kendler, T. S. Concept formation. Annual Review of Psychology, 1961, 12, 447-472.

Klausmeier, H. J., Wiersma, W., & Harris, C. W. Efficiency of initial learning and transfer by individuals, pairs, and quads. Journal of Educational Psychology, 1963, 54, 160-164.

Laughlin, P. R. Speed versus minimum-choice instructions in concept attainment. Journal of Experimental Psychology, 1964, 67, 596.

Laughlin, P. R. Selection strategies in concept attainment as a function of number of persons and stimulus display. Journal of Experimental Psychology, 1965, 70, 323-327.

Laughlin, P. R. Selection strategies in concept attainment as a function of number of relevant problem attributes. Journal of Experimental Psychology, 1966, 71, 773-776.

Laughlin, P. R. Focusing strategy for eight concept rules. Journal of Experimental Psychology, 1968, 77, 661-669.

Laughlin, P. R. Selection versus reception concept-attainment paradigms as a function of memory, concept rule, and concept universe. Journal of Educational Psychology, 1969, 60, 267-273.

Laughlin, P. R. Selection versus reception concept-attainment paradigms for individuals and cooperative pairs. Journal of Educational Psychology, 1972, 63, 116-122.

Laughlin, P. R. Selection strategies as normative and descriptive models: Comments on Johnson's "The validation of concept-learning strategies." Journal of Experimental Psychology: General, 1978, 107, 267-272.

Laughlin, P. R. & Doherty, M. A. Discussion versus memory in cooperative group concept attainment. Journal of Educational Psychology, 1967, 58, 123-128.

Laughlin, P. R. & Jordan, R. M. Selection strategies in conjunctive, disjunctive, and biconditional concept attainment. Journal of Experimental Psychology, 1967, 75, 188-193.

Laughlin, P. R., Kalowski, C. A., Metzler, M. E., Ostap, K. M., & Venclovas, S. M. Concept identification as a function of sensory modality, information, and number of persons. Journal of Experimental Psychology, 1968, 77, 335-340. (a)

Laughlin, P. R., McGlynn, R. P., Anderson, J. A., & Jacobson, E. S. Concept attainment by individuals versus cooperative pairs as a function of memory, sex, and concept rule. Journal of Personality and Social Psychology, 1968, 8, 410-417. (b)

Laughlin, P. R. & Sweeney, J. D. Individual-to-group and group-to-individual transfer in problem solving. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 246-254.

Lemke, E. A. & Hecht, J. T. Effects of degree of training, group size, and inductive ability on the tranfer of conceptual behavior. The Journal of Educational Research, 1971, 65, 43-45.

Lemke, E. A., Randle, K., & Robertshaw, C. W. Effects of degree of initial acquisition, group size, and general mental ability in concept learning and transfer. Journal of Educational Psychology, 1969, 60, 75-78.

Levine, M. A Cognitive Theory of Learning: Research on Hypothesis-Testing. Hillsdale, NJ: Lawrence Erlbaum Associates, 1975.

McGlynn, R. P. Four-person group concept attainment as a function of interaction format. The Journal of Social Psychology, 1972, 86, 89-94.

McGlynn, R. P. & Schick, C. Dyadic concept attainment as a function of interaction format, memory requirements, and sex. Journal of Educational Psychology, 1973, 65, 335-340. (a)

McGlynn, R. P. & Schick, C. Reduced information feedback on a selection côncept learning task. Behavioral Research Methods and Instrumentation, 1973, 5, 335-356. (b)

Miller, L. A. Hypothesis analysis of conjunctive concept-learning situations. Psychological Review, 1971, 78, 262-271.

Neisser, U. & Weene, P. Hierarchies in concept attainment. Journal of Experimental Psychology, 1962, 64, 640-645.

Newell, A., Shaw, J. C., & Simon, H. A. Elements of a theory of human problem solving. Psychological Review, 1958, 65, 151-166.

Osler, S. F. & Fivel, M. W. Concept attainment: I. The role of age and intelligence in concept attainment by induction. Journal of Experimental Psychology, 1961, 62, 1-8.

Osler, S. F. & Trautman, G. E. Concept attainment: II. Effects of stimulus complexity upon conept attainment at two levels of intelligence. Journal of Experimental Psychology, 1961, 62, 9-13.

Osler, S. F. & Weiss, S. R. Studies in concept attainment: III. Effect of instructions at two levels of intelligence. Journal of Experimental Psychology, 1962, 63, 528-533.

Royer, J. Theories of the transfer of learning. Educational Psychologist, 1979, 14, 53-60.

Schick, C. & McGlynn, R. P. Cooperation versus competition in group concept attainment under conditions of information exchange. The Journal of Social Psychology, 1976, 100, 311-312.

Smoke, K. L. Negative instances in concept learning. Journal of Experimental Psychology, 1933, 16, 583-588.

Tatsuoka, M. M. Multivariate analysis in educational research. Review of Research in Education, 1974, 1, 273-319.

Tumblin, A. & Gholson, B. Hypothesis theory and the development of conceptual learning. Psychological Bulletin, 1981, 90, 102-124.

Wells, H. Effects of transfer and problem structure in disjunctive concept formation. Journal of Experimental Psychology, 1963, 65, 63-69.

APPENDIX A. SAMPLE INSTRUCTIONS

The following instructions were used in Experiment 1, during the learning session. Similar instructions were used in the other experiments, with appropriate modifications. These instructions were for subjects working in pairs. Other than omitting all references to discussion, cooperation, and team effort, the instructions are the same as those delivered to individuals.

> On the table in front of you is an array of three different geometric shapes, of three different sizes, and three different colors. The shapes, sizes and colors are called values of the dimensions of shape, size and color. There are 27 different combinations of these values in this array. Each combination is called an instance of all possible combinations. It is possible to have rules which describe different sets of these instances. For example, the simple rule of 'red' would include all of the red instances on the array. Such a simple rule is called a single-value rule. There are more complicated rules possible. One such rule is called the two-value conjunctive rule. Conjunctive rules combine two values by use of the word 'and.' For an instance to be contained in a set described by a conjunctive rule, it must contain all of the values which form the rule. For example, the conjunctive rule of 'red and circle' would contain only those instances which are both red and a circle. Another type of rule is the two-value disjunctive rule. Disjunctive rules combine values by the use of the word 'or.' For example, 'red or circle' would include all red instances and any circle. For an instance to belong to a disjunctive rule it must contain at least one of the values of the rule. Any questions?
>
> I have a list of rules. Some of these may be single-value rules, some may be two-value conjunctives and some may be two-value disjunctives. I want you to try to identify each rule; not only the type of rule, but the exact rule itself. The procedure we will use is as follows. First,

I will describe for you an instance and tell you whether or not it is in the set described by the first rule. Then I want you to tell me a possible rule, that is, tell me a rule you think could be mine, on the basis of the instance I provide. After telling me a possible rule, you select an instance. I will tell you whether or not your selected instance is in the set described by the rule to be identified. I will not tell you if your possible rules are correct or not. We will continue this sequence of stating a possible rule--selecting an instance--receiving feedback until you think you are positive of the identity of the rule. If you are positive you know the rule, tell me, 'Our conclusion is. . .' and state the rule. If you are correct, we will start the next problem. If you are incorrect, I will have you make another selection of an instance and we will continue until you do solve the problem. Please make your conclusions only if you are certain you know the rule.

Since the two of you are working together, you may discuss anything you want. I want you to cooperate with each other. Please come to an agreement for each selected instance and possible rule before telling them to me. Remember, I want this to be a team effort. I will be recording your choices and timing you. Please do not be concerned with the time. I am more interested in what selections you use than in how fast you solve the problems. Any questions?

APPENDIX B. ANOVA SUMMARY TABLES

Table B1. ANOVA table for trials to solution during learning Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.0563	0.0563	0.00	0.9490	
Group Size	1	0.1563	0.1563	0.01	0.0151	
Sex x Group Size	1	23.2563	23.2563	1.72	0.1984	
Sub (Sex x Group Size)	36	487.7250	13.5479			
Problems	3	56.6188	18.8729	2.15	0.0969	
Sex x Problems	3	14.6188	4.8729	0.55	0.6502	
Group Size x Problems	3	43.5188	14.5063	1.65	0.1803	
Sex x Group Size x Problems	3	31.7188	10.5729	1.20	0.3119	
Sub (Sex x Group Size) x Problems	108	948.7750	8.7850			

Table B2. ANOVA table for untenable hypotheses during learning in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.4261	0.4261	0.32	0.5774	
Group Size	1	0.2412	0.2412	0.18	0.6748	
Sex x Group Size	1	0.9159	0.9159	0.68	0.4151	
Sub (Sex x Group Size)	36	48.5095	1.3475			
Problems	3	4.7606	1.5869	3.50	0.0180	0.03
Sex x Problems	3	1.6380	0.5460	1.20	0.3121	
Group Size x Problems	3	0.3004	0.1001	0.22	0.8811	
Sex x Group Size x Problems	3	2.5653	0.8551	1.88	0.1350	
Sub (Sex x Group Size) x Problems	108	49.0247	0.4539			

Table B3. ANOVA table for strategy efficiency during learning in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.7453	0.7453	0.65	0.4254	
Group Size	1	0.0602	0.0602	0.05	0.8200	
Sex x Group Size	1	1.4335	1.4335	1.25	0.2709	
Sub (Sex x Group Size)	36	41.2766	1.1466			
Problems	3	5.7602	1.9201	5.99	0.0009	0.05
Sex x Problems	3	1.0568	0.3523	1.10	0.3551	
Group Size x Problems	3	2.1299	0.7100	2.22	0.0890	
Sex x Group Size x Problems	3	2.2786	0.7595	2.37	0.0733	
Sub (Sex x Group Size) x Problems	108	34.5962	0.3203			

Table B4. ANOVA table for decision time during learning in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	297.025	297.025	0.41	0.5238	
Group Size	1	3783.025	3783.025	5.28	0.0275	0.05
Sex x Group Size	1	60.025	60.025	0.08	0.7740	
Sub (Sex x Group Size)	36	25804.800	716.800			
Problems	3	10136.125	3378.708	18.83	0.0001	
Sex x Problems	3	392.825	130.942	0.73	0.5398	0.16
Group Size x Problems	3	615.425	205.142	1.14	0.3353	
Sex x Group Size x Problems	3	765.325	255.108	1.42	0.2394	
Sub (Sex x Group Size) x Problems	108	19381.800	179.461			

Table B5. ANOVA table for trials to solution during transfer in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	14.400	14.400	1.36	0.2519	
Group Size	1	36.100	36.100	3.40	0.0734	
Sex x Group Size	1	0.625	0.625	0.06	0.8097	
Sub (Sex x Group Size)	36	382.250	10.618			
Problems	3	41.125	13.708	1.69	0.1712	
Sex x Problems	3	11.750	3.917	0.48	0.6983	
Group Size x Problems	3	3.550	1.183	0.15	0.9286	
Sex x Group Size x Problems	3	17.025	5.675	0.70	0.5571	
Sub (Sex x Group Size) x Problems	108	874.550	8.098			

Table B6. ANOVA table for untenable hypotheses during transfer in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.3656	0.3656	0.43	0.5154	
Group Size	1	13.8745	13.8745	16.38	0.0003	0.10
Sex x Group Size	1	0.6853	0.6853	0.81	0.3744	
Sub (Sex x Group Size)	36	30.4918	0.8470			
Problems	3	5.5457	1.8486	2.75	0.0454	0.03
Sex x Problems	3	1.3799	0.4600	0.68	0.5669	
Group Size x Problems	3	0.8994	0.2998	0.45	0.7241	
Sex x Group Size x Problems	3	1.1364	0.3788	0.56	0.6440	
Sub (Sex x Group Size) x Problems	108	72.5434	0.6717			

Table B7. ANOVA table for strategy efficiency during transfer in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.5425	0.5425	0.60	0.4452	
Group Size	1	8.1511	8.1511	8.95	0.0050	0.08
Sex x Group Size	1	0.0487	0.0487	0.05	0.8185	
Sub (Sex x Group Size)	36	32.7777	0.9105			
Problems	3	6.1074	2.0358	5.93	0.0010	0.06
Sex x Problems	3	1.2081	0.4027	1.17	0.3232	
Group Size x Problems	3	0.8490	0.2830	0.82	0.4856	
Sex x Group Size x Problems	3	1.2199	0.4066	1.19	0.3188	
Sub (Sex x Group Size) x Problems	108	37.0542	0.3431			

Table B8. ANOVA table for decision time during transfer in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	50.625	50.625	0.18	0.6748	
Group Size	1	27.225	27.225	0.10	0.7582	
Sex x Group Size	1	19.600	19.600	0.07	0.7939	
Sub (Sex x Group Size)	36	10184.650	282.907			
Problems	3	7434.150	2478.142	23.61	0.0001	0.23
Sex x Problems	3	742.425	247.475	2.36	0.0745	
Group Size x Problems	3	1455.125	485.042	4.62	0.0046	0.04
Sex x Group Size x Problems	3	207.750	69.250	0.66	0.5822	
Sub (Sex x Group Size) x Problems	108	11333.550	104.940			

Table B9. ANOVA table for trials to solution between sessions in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	11.250	11.250	0.17	0.6801	
Group Size	1	33.800	33.800	0.52	0.4759	
Sex x Group Size	1	31.250	31.250	0.48	0.4929	
Sub (Sex x Group Size)	36	2343.900	65.108			
Sessions	1	26.450	26.450	0.80	0.3785	
Sex x Sessions	1	7.200	7.200	0.22	0.6446	
Group Size x Sessions	1	48.050	48.050	1.44	0.2373	
Sex x Group Size x Sessions	1	64.800	64.800	1.99	0.1713	
Sub (Sex x Group Size) x Sessions	36	1197.500	33.264			

Table B10. ANOVA table for untenable hypotheses between sessions in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.1504	0.1504	0.36	0.5547	
Group Size	1	0.8364	0.8364	1.98	0.1683	
Sex x Group Size	1	0.2555	0.2555	0.60	0.4421	
Sub (Sex x Group Size)	36	15.2277	0.4229			
Sessions	1	2.7023	2.7023	11.97	0.0014	0.08
Sex x Sessions	1	0.2273	0.2273	1.01	0.3224	
Group Size x Sessions	1	1.4076	1.4078	6.23	0.0172	0.04
Sex x Group Size x Sessions	1	0.1098	0.1098	0.49	0.4901	
Sub (Sex x Group Size) x Sessions	36	8.1292	0.2258			

Table B11. ANOVA table for strategy efficiency between sessions in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	0.0136	0.0136	0.05	0.8331	
Group Size	1	0.6406	0.6406	2.12	0.1541	
Sex x Group Size	1	0.3095	0.3095	1.02	0.3184	
Sub (Sex x Group Size)	36	10.8818	0.3023			
Sessions	1	0.8532	0.8532	9.29	0.0043	0.05
Sex x Sessions	1	0.0283	0.0283	0.31	0.5821	
Group Size x Sessions	1	0.6249	0.6249	6.80	0.0132	0.03
Sex x Group Size x Sessions	1	0.0233	0.0233	0.25	0.6178	
Sub (Sex x Group Size) x Sessions	36	3.3072	0.0917			

Table B12. ANOVA table for decision time between sessions in Experiment 1

Source	df	SS	MS	F	p	ω^2
Sex	1	84.050	84.050	0.41	0.5273	
Group Size	1	561.800	561.800	2.77	0.1046	
Sex x Group Size	1	54.450	54.450	0.27	0.6074	
Sub (Sex x Group Size)	36	7297.500	202.708			
Sessions	1	11.250	11.250	0.17	0.6821	
Sex x Sessions	1	12.800	12.800	0.19	0.6622	
Group Size x Sessions	1	396.050	396.050	6.00	0.0193	0.03
Sex x Group Size x Sessions	1	3.200	3.200	0.05	0.8269	
Sub (Sex x Group Size) x Sessions	36	2374.700	65.964			

Table B13. ANOVA table for trials to solution during learning in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	14.4453	14.4453	0.56	0.4601	
Array	3	107.4609	35.8203	1.40	0.2678	
Sex x Array	3	137.3984	45.7995	1.79	0.1766	
Sub (Sex x Array)	24	615.1875	25.6328			
Problems	3	292.7734	97.5911	5.05	0.0033	0.08
Sex x Problems	3	92.5859	30.8620	1.60	0.1965	
Array x Problems	9	226.8828	25.2092	1.30	0.2499	
Sex x Array x Problems	9	197.5625	21.9939	1.14	0.3484	
Sub (Sex x Array) x Problems	72	1392.5625	19.3411			

Table B14. ANOVA table for untenable hypotheses during learning in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.0976	0.0976	0.05	0.8294	
Array	3	1.0976	0.3659	0.18	0.9103	
Sex x Array	3	5.2418	1.7473	0.85	0.4805	
Sub (Sex x Array)	24	49.3678	2.0570			
Problems	3	3.0809	1.0270	1.95	0.1283	
Sex x Problems	3	0.9809	0.3270	0.62	0.6085	
Array x Problems	9	8.9224	0.9914	1.88	0.0687	
Sex x Array x Problems	9	7.8486	0.8721	1.65	0.1164	
Sub (Sex x Array) x Problems	72	38.0035	0.5278			

Table B15. ANOVA table for strategy efficiency during learning in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.2168	0.2168	0.19	0.6644	
Array	3	1.2620	0.4207	0.37	0.7721	
Sex x Array	3	2.9897	0.9966	0.89	0.4618	
Sub (Sex x Array)	24	26.9566	1.1231			
Problems	3	2.4114	0.8038	1.77	0.1591	
Sex x Problems	3	0.9601	0.3200	0.70	0.5559	
Array x Problems	9	2.5778	0.2864	0.63	0.7684	
Sex x Array x Problems	9	3.4819	0.3869	0.85	0.5727	
Sub (Sex x Array) x Problems	72	32.7075	0.4543			

Table B16. ANOVA table for decision time during learning in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	31.0078	31.0078	0.10	0.7601	
Array	3	3032.3359	1010.7786	3.11	0.0453	0.05
Sex x Array	3	517.1484	172.3828	0.53	0.6659	
Sub (Sex x Array)	24	7803.9375	325.1641			
Problems	3	11829.2734	3943.0910	19.78	0.0001	0.26
Sex x Problems	3	7.4609	2.4870	0.01	0.9946	
Array x Problems	9	4486.1953	498.4661	2.50	0.0151	0.06
Sex x Array x Problems	9	1604.5078	178.2786	0.89	0.5356	
Sub (Sex x Array) x Problems	72	14353.8125	199.3585			

Table B17. ANOVA table for trials to solution during transfer in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.3828	0.3828	0.03	0.8538	
Array	3	5.2734	1.7578	0.16	0.9226	
Sex x Array	3	14.1484	4.7161	0.43	0.7350	
Sub (Sex x Array)	24	264.6875	11.0286			
Problems	3	7.2109	2.4036	0.73	0.5389	
Sex x Problems	3	22.0859	7.3620	2.25	0.0890	
Array x Problems	9	26.8828	2.9870	0.91	0.5213	
Sex x Array x Problems	9	29.5078	3.2786	1.00	0.4484	
Sub (Sex x Array) x Problems	72	236.0625	3.2786			

Table B18. ANOVA table for untenable hypotheses during transfer in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	1.8751	1.8751	0.98	0.3333	
Array	3	6.9617	2.3206	1.21	0.3286	
Sex x Array	3	3.4959	1.1653	0.61	0.6175	
Sub (Sex x Array)	24	46.1510	1.9230			
Problems	3	0.8394	0.2798	0.71	0.5503	
Sex x Problems	3	2.0651	0.6884	1.76	0.1617	
Array x Problems	9	4.3490	0.4832	1.23	0.2888	
Sex x Array x Problems	9	5.7970	0.6441	1.64	0.1188	
Sub (Sex x Array) x Problems	72	28.2246	0.3920			

Table B19. ANOVA table for strategy efficiency during transfer in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.0197	0.0197	0.02	0.8800	
Array	3	7.4450	2.4817	2.93	0.0539	
Sex x Array	3	1.4111	0.4704	0.56	0.6492	
Sub (Sex x Array)	24	20.3085	0.8462			
Problems	3	0.5636	0.1879	0.69	0.5662	
Sex x Problems	3	0.2660	0.0887	0.32	0.8096	
Array x Problems	9	3.7791	0.4199	1.54	0.1512	
Sex x Array x Problems	9	4.8870	0.5430	1.99	0.0530	
Sub (Sex x Array) x Problems	72	19.6781	0.2733			

Table B20. ANOVA table for decision time during transfer in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	159.7578	159.7578	0.51	0.4811	
Array	3	3197.3984	1065.7995	3.42	0.0335	0.12
Sex x Array	3	241.7109	80.5730	0.26	0.8546	
Sub (Sex x Array)	24	7485.3125	311.8880			
Problems	3	520.0234	173.3411	2.35	0.0782	
Sex x Problems	3	184.8359	61.6118	0.84	0.4812	
Array x Problems	9	497.1328	55.2370	0.75	0.6640	
Sex x Array x Problems	9	571.3203	63.4800	0.86	0.5645	
Sub (Sex x Array) x Problems	72	5308.4375	73.7283			

Table B21. ANOVA table for trials to solution between sessions in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	20.2500	20.2500	0.21	0.6472	
Order	1	64.0000	64.0000	0.68	0.4178	
Sex x Order	1	361.0000	361.0000	3.81	0.0609	
Sub (Sex x Order)	28	2649.7500	99.6339			
Sessions	1	770.0625	770.0625	18.15	0.0002	0.13
Sex x Sessions	1	39.0625	39.0625	0.92	0.3456	
Order x Sessions	1	175.0625	175.0625	4.14	0.0515	
Sex x Order x Sessions	1	138.0625	138.0625	3.25	0.0820	
Sub (Sex x Order) x Sessions	28	1188.2500	42.4375			

Table B22. ANOVA table for untenable hypotheses between sessions in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.3584	0.3584	0.53	0.4738	
Order	1	0.0526	0.0526	0.08	0.7830	
Sex x Order	1	2.4225	2.4225	3.56	0.0695	
Sub (Sex x Order)	28	19.0360	0.6799			
Sessions	1	1.1701	1.1701	6.03	0.0205	0.03
Sex x Sessions	1	0.0469	0.0469	0.24	0.6266	
Order x Sessions	1	0.2634	0.2634	0.36	0.2537	
Sex x Order x Sessions	1	0.1313	0.1313	0.68	0.4176	
Sub (Sex x Order) x Sessions	28	5.4325	0.1940			

Table B23. ANOVA table for strategy efficiency between sessions in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	0.4692	0.4692	2.00	0.1685	
Order	1	0.1270	0.1270	0.54	0.4681	
Sex x Order	1	0.4014	0.4014	1.71	0.2016	
Sub (Sex x Order)	28	6.5738	0.2348			
Sessions	1	0.2545	0.2545	2.22	0.1477	
Sex x Sessions	1	0.0000	0.0000	0.00	0.9903	
Order x Sessions	1	0.1015	0.1015	0.88	0.3551	
Sex x Order x Sessions	1	0.2658	0.2658	2.32	0.1393	
Sub (Sex x Order) x Sessions	28	3.2149	0.1148			

Table B24. ANOVA table for decision time between sessions in Experiment 2

Source	df	SS	MS	F	p	ω^2
Sex	1	3.0625	3.0625	0.01	0.9104	
Order	1	115.5625	115.5625	0.49	0.4910	
Sex x Order	1	272.2500	272.2500	1.15	0.2933	
Sub (Sex x Order)	28	6644.8750	237.3170			
Sessions	1	1040.0625	1040.0625	13.99	0.0008	0.09
Sex x Sessions	1	90.2500	90.2500	1.21	0.2798	
Order x Sessions	1	30.2500	30.2500	0.41	0.5287	
Sex x Order x Sessions	1	22.5625	22.5625	0.30	0.5860	
Sub (Sex x Order) x Sessions	28	2080.8750	74.3170			

Table B25. ANOVA table for trials to solution during learning in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	46.4101	46.4101	6.72	0.0126	0.03
Vocalization	1	8.6289	8.6289	1.25	0.2692	
Array	3	84.4180	28.1393	4.08	0.0117	0.05
Groups x Vocalization	1	11.8164	28.1393	1.71	0.1970	
Groups x Array	3	2.1055	0.7018	0.10	0.9587	
Vocalization x Array	3	56.7617	18.9206	2.74	0.0535	
Groups x Vocalization x Array	3	22.4492	7.4831	1.08	0.3650	
Sub (Groups x Vocalization x Array	48	331.4375	6.9049			
Problems	3	125.7617	41.9206	12.55	0.0001	0.08
Groups x Problems	3	4.6367	1.5456	0.46	0.7125	
Vocalization x Problems	3	23.9805	7.9935	2.39	0.0697	
Array x Problems	9	95.7852	10.6428	3.19	0.0016	0.05
Groups x Vocalization x Problems	3	19.1055	6.3685	1.91	0.1293	
Groups x Array x Problems	9	17.5352	1.9484	0.58	0.8101	
Vocalization x Array x Problems	9	50.5664	5.6185	1.68	0.0977	
Groups x Vocalization x Array x Problems	9	31.5663	3.5074	1.05	0.4034	
Sub (Groups x Vocalization x Array) x Problems	144	480.8125	3.3390			

Table B26. ANOVA table for untenable hypotheses during learning in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	3.7775	3.7775	5.39	0.0246	0.02
Vocalization	1	4.5247	4.5247	6.45	0.0144	0.03
Array	3	13.7370	4.5790	6.53	0.0009	0.09
Groups x Vocalization	1	4.1142	4.1142	5.87	0.0192	0.03
Groups x Array	3	0.9556	0.3185	0.45	0.7154	
Vocalization x Array	3	0.9680	0.3227	0.46	0.7113	
Groups x Vocalization x Array	3	2.9632	0.9877	1.41	0.2516	
Sub (Groups x Vocalization x Array)	48	33.6461	0.7010			
Problems	3	2.0544	0.6848	2.18	0.0913	
Groups x Problems	3	0.5867	0.1956	0.62	0.6052	
Vocalization x Problems	3	0.5367	0.1789	0.57	0.6399	
Array x Problems	9	3.8019	0.4224	1.35	0.2180	
Groups x Vocalization x Problems	3	1.7035	0.5678	1.81	0.1465	
Groups x Array x Problems	9	2.6709	0.2968	0.95	0.4887	
Vocalization x Array x Problems	9	3.7331	0.4148	1.32	0.2302	
Groups x Vocalization x Array x Problems	9	2.4834	0.2759	0.88	0.5465	
Sub (Groups x Vocalization x Array) x Problems	144	45.2070	0.3139			

Table B27. ANOVA table for strategy efficiency during learning in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	5.7897	5.7897	15.90	0.0002	0.07
Vocalization	1	0.8505	0.8505	2.34	0.1330	
Array	3	6.6615	2.2205	6.10	0.0013	0.07
Groups x Vocalization	1	1.9011	1.9011	5.22	0.0268	0.02
Groups x Array	3	1.3137	0.4379	1.20	0.3189	
Vocalization x Array	3	0.3440	0.1147	0.31	0.8145	
Groups x Vocalization x Array	3	1.0022	0.3340	0.92	0.4396	
Sub (Groups x Vocalization x Array)	48	17.4795	0.3642			
Problems	3	4.8285	1.6095	7.12	0.0002	0.05
Groups x Problems	3	0.6049	0.2016	0.89	0.4491	
Vocalization x Problems	3	0.0442	0.0147	0.07	0.9727	
Array x Problems	9	1.0247	0.1139	0.50	0.8707	
Groups x Vocalization x Problems	3	1.3491	0.4497	0.99	0.1167	
Groups x Array x Problems	9	0.7404	0.0823	0.36	0.9500	
Vocalization x Array x Problems	9	3.5185	0.3909	1.73	0.0870	
Groups x Vocalization x Array x Problems	9	1.3632	0.1515	0.67	0.7364	
Sub (Groups x Vocalization x Array) x Problems	144	32.5608	0.2261			

Table B28. ANOVA table for decision time during learning in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	7003.5977	7003.5977	14.60	0.0004	0.06
Vocalization	1	8021.4414	8021.4414	16.73	0.0002	0.07
Array	3	3024.6054	1068.2018	2.23	0.0970	
Groups x Vocalization	1	119.6289	119.6289	0.25	0.6197	
Groups x Array	3	1851.8242	617.2747	1.29	0.2895	
Vocalization x Array	3	635.3555	211.7852	0.44	0.7243	
Groups x Vocalization x Array	3	2108.5430	702.8476	1.47	0.2357	
Sub (Groups x Vocalization x Array)	48	23018.5625	479.5534			
Problems	3	28315.7930	9438.5973	72.28	0.0001	0.27
Groups x Problems	3	497.2617	165.7539	1.27	0.2865	
Vocalization x Problems	3	374.3555	124.7852	0.96	0.4171	
Array x Problems	9	2136.0977	237.3442	1.82	0.0694	
Groups x Vocalization x Problems	3	749.2930	249.7643	1.91	0.1284	
Groups x Array x Problems	9	3201.7539	355.7504	2.72	0.0059	0.006
Vocalization x Array x Problems	9	486.2852	54.0317	0.41	0.9260	
Groups x Vocalization x Array x Problems	9	2536.2227	281.8025	2.16	0.0280	-*
Sub (Groups x Vocalization x Array x Problems	144	18803.1875	130.5777			

*Less than 0.005.

Table B29. ANOVA table for trials to solution during transfer in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	1.5625	1.5625	0.42	0.5197	
Vocalization	1	1.5625	1.5625	0.42	0.5197	
Array	3	38.8281	12.9427	3.49	0.0227	0.04
Groups x Vocalization	1	1.2656	1.2656	0.34	0.5621	
Groups x Array	3	14.7188	4.9063	1.32	0.2784	
Vocalization x Array	3	12.7813	4.2604	1.15	0.3396	
Groups x Vocalization x Array	3	9.6406	3.2135	0.87	0.4656	
Sub (Groups x Vocalization x Array)	48	178.2500	3.7135			
Problems	3	37.3906	12.4635	7.71	0.0001	0.05
Groups x Problems	3	3.6563	1.2188	0.75	0.5249	
Vocalization x Problems	3	20.6463	6.884	4.26	0.0066	0.02
Array x Problems	9	36.1406	4.0156	2.48	0.0115	0.03
Groups x Vocalization x Problems	3	1.7656	0.5885	0.36	0.7818	
Groups x Array x Problems	9	7.3124	0.8125	0.50	0.8713	
Vocalization x Array x Problems	9	23.0000	2.5556	1.58	0.1255	
Groups x Vocalization x Array x Problems	9	16.3281	1.8142	1.12	0.3505	
Sub (Groups x Vocalization x Array x Problems	144	232.7500	1.6163			

Table B30. ANOVA table for untenable hypotheses during transfer in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	0.3049	0.3049	0.29	0.5918	
Vocalization	1	15.6602	15.6602	14.96	0.0003	0.10
Array	3	7.1788	2.3929	2.29	0.0906	
Groups x Vocalization	1	0.0411	0.0411	0.04	0.8438	
Groups x Array	3	9.1889	3.0630	2.93	0.0431	0.04
Vocalization x Array	3	0.8400	0.2800	0.27	0.8485	
Groups x Vocalization x Array	3	1.2685	0.4228	0.40	0.7408	
Sub (Groups x Vocalization x Array)	48	50.2359	1.0466			
Problems	3	3.8930	1.2977	4.59	0.0044	0.02
Groups x Problems	3	1.0270	0.3423	1.21	0.3076	
Vocalization x Problems	3	0.0824	0.0824	0.29	0.8327	
Array x Problems	9	3.6522	0.4058	1.44	0.1773	
Groups x Vocalization x Problems	3	0.3525	0.1175	0.42	0.7455	
Groups x Array x Problems	9	1.8691	0.2077	0.73	0.6777	
Vocalization x Array x Problems	9	2.1079	0.2342	0.83	0.5918	
Groups x Vocalization x Array x Problems	9	2.1845	0.2427	0.86	0.5645	
Sub (Groups x Vocalization x Array) x Problems	144	40.6963	0.2826			

Table B31. ANOVA table for strategy efficiency during transfer in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	0.1365	0.1365	0.34	0.5619	
Vocalization	1	1.7102	1.7102	4.28	0.0441	0.02
Array	3	1.6026	0.5342	1.34	0.2738	
Groups x Vocalization	1	0.0001	0.0001	0.00	0.9856	
Groups x Array	3	1.4317	0.4772	1.19	0.3223	
Vocalization x Array	3	1.7598	0.5866	1.47	0.2354	
Groups x Vocalization x Array	3	0.1835	0.0612	0.15	0.9273	
Sub (Groups x Vocalization x Array)	48	19.1972	0.3999			
Problems	3	2.0424	0.6808	5.39	0.0016	0.03
Groups x Problems	3	0.2569	0.0856	0.68	0.5702	
Vocalization x Problems	3	0.0931	0.0310	0.25	0.8642	
Array x Problems	9	2.8910	0.3213	2.55	0.0097	0.03
Groups x Vocalization x Problems	3	0.5921	0.1974	1.56	0.1992	
Groups x Array x Problems	9	1.5774	0.1753	1.39	0.1977	
Vocalization x Array x Problems	9	1.6295	0.1811	1.43	0.1779	
Groups x Vocalization x Array x Problems	9	0.6886	0.0765	0.61	0.7913	
Sub (Groups x Vocalization x Array) x Problems	144	18.1733	0.1262			

Table B32. ANOVA table for decision times during transfer in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	4.7852	4.7852	0.02	0.8973	
Vocalization	1	170.6289	170.6289	0.60	0.4423	
Array	3	5739.9805	1913.3268	6.73	0.0007	0.11
Groups x Vocalization	1	65.0039	65.0039	0.23	0.6347	
Groups x Array	3	631.8867	210.6289	0.74	0.5330	
Vocalization x Array	3	560.6680	186.8893	0.66	0.5823	
Groups x Vocalization x Array	3	318.6055	106.2018	0.37	0.7725	
Sub (Groups x Vocalization x Array)	48	13647.3125	284.3190			
Problems	3	3648.6055	1216.2018	10.44	0.0001	0.07
Groups x Problems	3	435.8867	145.2956	1.25	0.2943	
Vocalization x Problems	3	182.6680	60.8893	0.52	0.6713	
Array x Problems	9	846.8466	94.0942	0.81	0.6106	
Groups x Vocalization x Problems	3	99.3555	33.1185	0.28	0.8377	
Groups x Array x Problems	9	1152.0039	128.0004	1.10	0.3670	
Vocalization x Array x Problems	9	828.0977	92.0109	0.79	0.6270	
Groups x Vocalization x Array x Problems	9	606.3477	67.3720	0.58	0.8141	
Sub (Groups x Vocalization x Array) x Problems	144	16767.9375	166.4440			

Table B33. ANOVA table for trials to solution between sessions in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	126.0078	126.0078	4.61	0.0361	0.03
Vocalization	1	51.2578	51.2578	1.88	0.1762	
Order	1	9.5703	9.5703	0.35	0.5563	
Groups x Vocalization	1	20.3203	20.3203	0.74	0.3921	
Groups x Order	1	0.3828	0.3828	0.01	0.9062	
Vocalization x Order	1	207.5703	207.5703	7.60	0.0079	0.05
Groups x Vocalization x Order	1	51.2578	51.2578	1.88	0.1762	
Sub (Groups x Vocalization x Order)	56	1529.5625	27.3136			
Session	1	61.8828	61.8828	3.28	0.0754	
Groups x Session	1	82.8828	82.8828	4.40	0.0406	0.01
Vocalization x Session	1	4.8828	4.8828	0.26	0.6129	
Order x Session	1	321.9453	321.9453	17.07	0.0001	0.08
Groups x Vocalization x Session	1	39.3828	39.2838	2.09	0.1540	
Groups x Order x Session	1	15.8203	15.8203	0.84	0.3636	
Vocalization x Order x Session	1	39.3828	39.3828	2.09	0.1540	
Groups x Vocalization x Order x Session	1	51.2578	51.2578	2.72	0.1048	
Sub (Groups x Vocalization x Order) x Session	56	1056.0625	18.8583			

Table B34. ANOVA table for untenable hypotheses between sessions in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	0.7293	0.7293	1.77	0.1894	
Vocalization	1	5.2122	5.2122	12.62	0.0008	0.11
Order	1	1.2582	1.2582	3.05	0.0864	
Groups x Vocalization	1	0.4403	0.4403	1.07	0.3064	
Groups x Order	1	0.4604	0.4604	1.11	0.2957	
Vocalization x Order	1	0.0104	0.0104	0.03	0.8743	
Groups x Vocalization x Order	1	0.4022	0.4022	0.97	0.3281	
Sub (Groups x Vocalization x Order)	56	23.1365	0.4132			
Session	1	1.0256	1.0256	7.97	0.0066	0.02
Groups x Session	1	0.9706	0.9706	7.54	0.0081	0.02
Vocalization x Session	1	0.2285	0.2285	1.78	0.1881	
Order x Session	1	3.1512	3.1512	24.48	0.0001	0.07
Groups x Vocalization x Session	1	0.3423	0.3423	2.66	0.1086	
Groups x Order x Session	1	0.0017	0.0017	0.01	0.9083	
Vocalization x Order x Session	1	0.0488	0.0488	0.38	0.5406	
Groups x Vocalization x Order x Session	1	0.0206	0.0206	0.16	0.6907	
Sub (Groups x Vocalization x Order) x Session	56	7.2092	0.1287			

Table B35. ANOVA table for strategy efficiency between sessions in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	0.4995	0.4995	4.07	0.0485	0.03
Vocalization	1	0.7818	0.7818	6.37	0.0145	0.05
Order	1	0.4610	0.4610	3.76	0.0577	
Groups x Vocalization	1	0.1440	0.1440	1.17	0.2833	
Groups x Order	1	0.1423	0.1423	1.16	0.2862	
Vocalization x Order	1	0.0141	0.0141	0.11	0.7363	
Groups x Vocalization x Order	1	0.0007	0.0007	0.01	0.9420	
Sub (Groups x Vocalization x Order)	56	6.8728	0.1227			
Session	1	0.1583	0.1583	4.87	0.0314	0.01
Groups x Session	1	0.7762	0.7762	23.89	0.0001	0.06
Vocalization x Session	1	0.0151	0.0151	0.47	0.4976	
Order x Session	1	0.9957	0.9957	30.64	0.0001	0.07
Groups x Vocalization x Session	1	0.1553	0.1553	4.78	0.0330	0.01
Groups x Order x Session	1	0.0605	0.0605	1.86	0.1778	
Vocalization x Order x Session	1	0.0161	0.0161	0.50	0.4846	
Groups x Vocalization x Order x Session	1	0.0919	0.0919	2.83	0.0982	
Sub (Groups x Vocalization x Order) x Session	56	1.8199	0.3250			

Table B36. ANOVA table for decision time between sessions in Experiment 3

Source	df	SS	MS	F	p	ω^2
Groups	1	825.1953	825.1953	6.00	0.0174	0.03
Vocalization	1	1478.3203	1478.3203	10.75	0.0018	0.06
Order	1	7.5078	7.5078	0.05	0.8161	
Groups x Vocalization	1	114.3828	114.3828	0.83	0.3656	
Groups x Order	1	31.0078	31.0078	0.23	0.6367	
Vocalization x Order	1	14.4453	14.4453	0.11	0.7470	
Groups x Vocalization x Order	1	122.0703	122.0703	0.89	0.3501	
Sub (Groups x Vocalization x Order)	56	7699.5625	137.4975			
Session	1	5447.0703	5447.0703	100.40	0.0001	0.25
Groups x Session	1	765.3828	765.3828	14.11	0.0004	0.03
Vocalization x Session	1	736.3203	736.3203	13.57	0.0005	0.03
Order x Session	1	689.1328	689.1328	12.70	0.0008	0.03
Groups x Vocalization x Session	1	2.2578	2.2578	0.40	0.8391	
Groups x Order x Session	1	239.2578	239.2578	4.41	0.0402	0.01
Vocalization x Order x Session	1	99.7578	99.7578	1.84	0.1806	
Groups x Vocalization x Order x Session	1	0.0078	0.0078	0.00	0.9905	
Sub (Groups x Vocalization x Order) x Session	56	3038.3125	54.2556			

APPENDIX C. THREE- AND FOUR-WAY INTERACTIONS

Table C1. Mean decision times for group size x arrays x problems in the learning session of Experiment 3[a]

Group	Problem 1	2	3	4
Individuals				
Forms	59.750	28.750	26.625	22.875
Cars	46.750	33.500	30.500	27.125
Borders	46.375	30.625	26.125	25.000
Houses	43.750	29.125	26.500	28.500
Pairs				
Forms	58.250	44.000	35.375	33.375
Cars	85.250	49.500	41.500	37.000
Borders	55.750	35.875	26.750	26.125
Houses	53.500	36.750	48.375	31.875

[a]HSD = 22.826, $p < 0.01$.

Table C2. Mean decision times for group size x vocalization x arrays x problems in the learning session of Experiment 3[a]

Group	Problem							
	Individuals				Pairs			
	1	2	3	4	1	2	3	4
Non-vocalizing								
Forms	64.75	27.25	23.50	17.25	47.50	31.50	26.75	28.50
Cars	32.00	27.24	24.74	23.25	96.50	39.75	41.00	36.75
Borders	36.75	28.00	21.24	21.75	49.75	29.00	20.50	21.50
Houses	37.50	19.25	25.50	22.25	46.00	25.75	32.50	25.50
Vocalizing								
Forms	54.75	30.25	29.75	28.50	69.00	56.50	44.00	28.35
Cars	61.50	39.75	35.25	31.00	74.00	59.24	42.00	37.25
Borders	56.00	33.24	31.00	28.25	61.75	42.75	33.00	30.75
Houses	50.00	39.00	27.50	34.75	61.00	47.75	64.25	38.25

[a]HSD = 32.281, $p < 0.01$.

Table C3. Mean decision times for group size x grouped arrays x session in Experiment 3

Group	Session: Learning	Session: Transfer
Individuals		
Perceptual-to-Semantic	33.813[a]	27.563
Semantic-to-Perceptual	34.250[b]	24.188
Pairs		
Perceptual-to-Semantic	40.063[c]	29.500
Semantic-to-Perceptual	47.938[d]	22.625

[a]Significantly more than Pairs/Semantic-to-Perceptual/Transfer at $p < 0.01$, and more than Individuals/Semantic-to-Perceptual/Transfer at $p < 0.05$.

[b]Significantly more than Semantic-to-Perceptual/Transfer of either group size at $p < 0.01$.

[c]Significantly more than any Transfer score at $p < 0.01$.

[d]Significantly more than any Individuals score and any Pairs/Transfer score at $p < 0.01$.